DOCUMENT RESUME

ED 092 592 95 TM 003 704

TITLE	Progress in Evaluation Study. Third Annual Report of the Center for the Study of Evaluation of Instructional Programs.
INSTITUTION	California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
PUB DATE	1 Nov 68
CONTRACT	OEC-4-6-061646-1909
NOTE	180p.
EDRS PRICE	MF-$0.75 HC-$9.00 PLUS POSTAGE
DESCRIPTORS	*Annual Reports; Evaluation; *Program Descriptions; *Research and Development Centers
IDENTIFIERS	*Center for the Study of Evaluation

ABSTRACT

This report covers the activities of the Center for the Study of Evaluation (CSE) from October 1, 1967 through October 31, 1968, and indicates the major efforts projected for the future. Included in the report are descriptions of the Center's projects and activities, the new CSE approach, administrative functions and supporting services, its staff and national advisors. (MLP)

U S DEPARTMENT OF HEALTH.
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM

ED 092592

PROGRESS
IN
EVALUATION
STUDY

TM 003 704

**CENTER FOR THE
STUDY OF
EVALUATION**
of instructional programs

EDUCATIONAL
CSE
RESEARCH AND DEVELOPMENT

3rd
Annual Report
to the
U. S. Office
of
Education

November 1, 1968

UCLA
Graduate School of Education
Los Angeles, California

ED 092592

PROGRESS IN EVALUATION STUDY

THIRD ANNUAL REPORT
OF THE
CENTER FOR THE STUDY OF EVALUATION
OF INSTRUCTIONAL PROGRAMS

November 1, 1968

UCLA
Graduate School of Education
Los Angeles, California

U. S. Office of Education
Government Contract:
OEC 4-6-061646-1909

UNIVERSITY OF CALIFORNIA, LOS ANGELES

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

GRADUATE SCHOOL OF EDUCATION
CENTER FOR THE STUDY OF EVALUATION
OF INSTRUCTIONAL PROGRAMS
LOS ANGELES, CALIFORNIA 90024

October 31, 1968

Division of Educational Laboratories
Bureau of Research
United States Office of Education
Washington, D. C. 20202

Gentlemen:

We are pleased to submit this Third Annual Report of the Center for the Study of Evaluation. The report covers CSE activities from October 1, 1967, through October 31, 1968, and indicates the major efforts projected for the future.

During the report period, CSE (the title has been shortened for easier identification from Center for the Study of Evaluation of Instructional Programs) made significant progress toward its goal of developing new theories and methods of evaluation. A wide range of studies were conducted, and many papers have been published for use by the academic community. Most importantly, the Center was able to develop a new organizational and programatic approach that will carry it more effectively toward its goal.

This report describes the new CSE approach, the continuing study projects included in it, the work produced during the period, and the support activities now being emphasized.

The reorganization and reorientation of CSE add to our expectations of accelerated progress in the coming year.

Sincerely yours,

Marvin C. Alkin

Marvin C. Alkin
Director

TABLE OF CONTENTS

SECTION PAGE

1. Introduction: A New Approach........................ 1

2. Terminating Projects

Foreword... 8

Instructional Variables Program........................ 9

Contextual Variables Program........................ 14

Criterion Measures Program........................ 44

3. New and Continuing Projects

Foreword... 62

Program on the Evaluation of Instructional Programs........................ 68

Program on the Evaluation of Educational Systems........................ 81

Program on Evaluation Theory and Methodology........................ 93

4. Administration

Reorganization and Redirection........................118

5. Supporting Services

Statistical..129

Dissemination......................................133

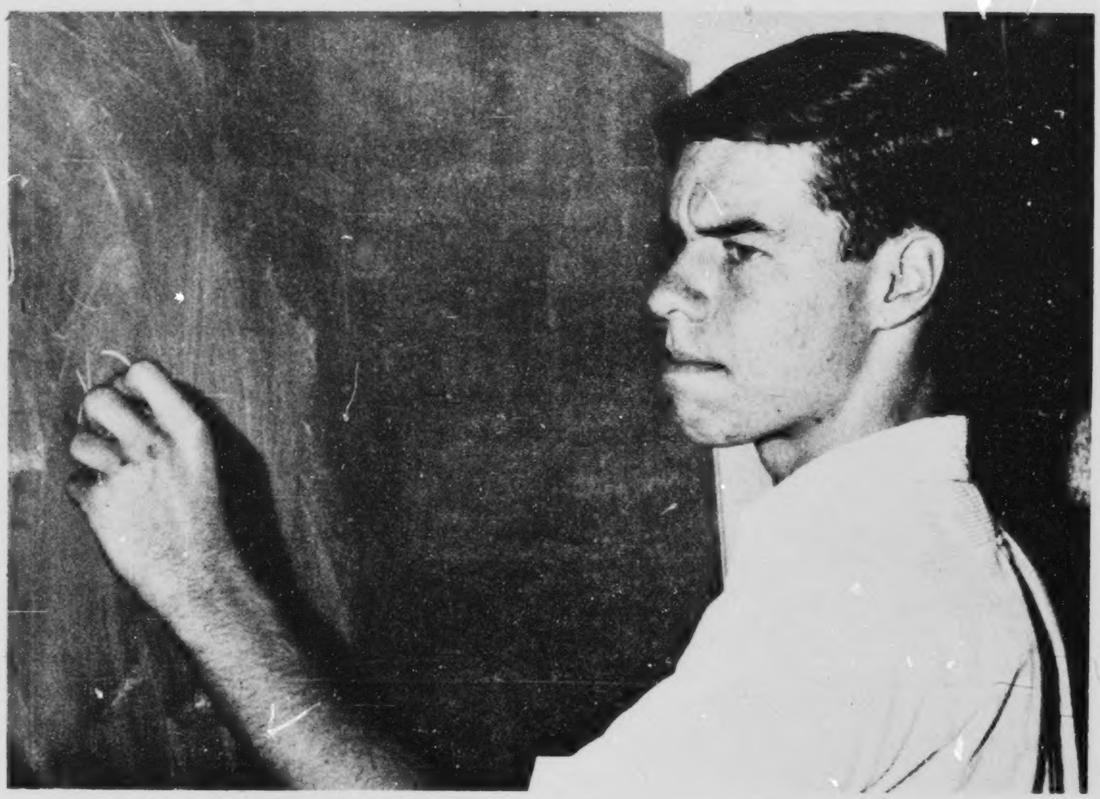
Clerical...142

6. The Center's Staff and National Advisors..............143

7. A Summary of Progress..............................155

8. Bibliography..166

INTRODUCTION



A NEW APPROACH

Since its inception in June, 1966, CSE has argued that evaluations of educational programs and systems have been inadequate because they have not dealt with a sufficient range of criterion, contextual, and instrumental variables. The Center expects to document this argument by showing how one's conclusions are changed by different combinations of data, by showing the amount of distortion that results from an inadequate range of variables, and by showing the types of variables that are most crucial for a balanced assessment.

Out of such study, new conceptualizations and models can be generated and tested. The point is to influence how people think about evaluation. More complex models are needed, both in the evaluation of instructional programs and in the evaluation of educational systems.

For the most part, the evaluation of instructional programs has concerned itself with rather direct, immediate, and specifically observable situations--such as studies of learning and teaching and curricula, of classrooms, local schools, individual performance, teacher-pupil interactions, etc.

FOCUS ON INTERACTIONS

The intent of such evaluation has been to promote clarification, focus, immediate feedback, and the continuous improvement of treatments and practices. More practitioners are engaged in work at this level than at any other; and much of their work has achieved a moderate level of sophistication--at least in the sense that people seem to know what they are doing, and how to go about it.

The basic weakness is that the work tends to deal with one problem at a time, rather than with more complex interactions. The special CSE contribution here is its concentration on the complex interactions which must be understood in order to evaluate adequately what goes on in the naturalistic school setting.

At the level of evaluation concerned with schools and school systems, the common practice has been one of describing and monitoring and plotting changes in the operation. Feedback for administrative or supervisory decisions is usually intended. The need here is to make such analysis and reports with more meaningful data than have typically been used.

To chart the impact of innovations or other modifications that, in varying degrees, work their way through the system, encountering various conditions that obstruct or facilitate their adoption or impact, certain kinds of observations, measures, and interactions must be studied. One is concerned with "systems changes" and the consequences of such changes. Evaluation at this level has been lacking sophistication. New ideas about what to look for and how to find it--and, of course, what to do with it analytically and conceptually--are urgently needed. Important contributions are especially needed here in view of legislative mandates to "evaluate" system level programs and special new programs.

Another important area of Center concern deals with the development and implementation of new data gathering and data analysis techniques which would be of value in improving the technology of evaluation.

CSE OBJECTIVES

The objectives of the Center can be defined, briefly, as follows:

1. To develop concepts of organization relevant to different levels of educational phenomena.
2. To develop working models of these concepts that will serve to: (a) define what kinds or variables and relationships should be studied, and (b) describe how the results of these evaluation studies can provide valid information for policy and program decisions.
3. To develop methods and methodologies needed in the practical conduct of evaluation studies, such as: (a) the integration of measuring instruments, experimental techniques, and statistical procedures, and (b) the development of new kinds of instruments and other data gathering methods.

STRENGTHENING PROGRAMMATIC RESEARCH

In its first two years, the Center has devoted a considerable amount of its research efforts to an examination of variables relevant for total evaluation models. Thus, the organizational structure included programs in instructional variables, individual contexts, organizational contexts, and criteria. CSE considered this a necessary phase in its development because of the paucity of educational research capital in this area. (See Francis S. Chase's recent article in the Journal of Research and Development in Education.) The Center's exploratory work has provided sufficient insights that it is now felt that maximum productivity can be achieved by concentrating to a greater extent on developing evaluation models and systems.

The newly strengthened programmatic thrust of the Center in no way implies a lack of concern for basic research activities. To the contrary, CSE feels that it is of utmost importance that selective research activities be sustained at a high level in order to broaden the knowledge base and provide the conceptual resources for what will ultimately become the evaluation models and systems referred to throughout this document.

In the process of developing evaluation models and systems, it is anticipated that voids will be found in the research related to specific potential elements of larger systems. Research can then be directed to filling these gaps within the context of the larger developmental and conceptual studies.

INTEGRATING SKILLS AND IDEAS

Within the scope of the programmatic approach, the activities of the Center become necessarily more interdisciplinary. They become more fully integrated and draw upon the skills and knowledge of a variety of research experts. The entire program is informed by the continuing modificaton of materials and procedures to meet the evolving demands of evaluation.

With this programmatic approach CSE moves toward greater concern with the study of evaluation itself, rather than the conduct of individual research studies on elements of a tentative evaluation model. The approach, then, is more integrated and interdisciplinary. There is a constant seeking after the nature of evaluation itself, and increased attention to total evaluation models.

CSE has organized its staff into smoothly functioning research teams which, through the programmatic approach, allows each to make a more significant personal contribution to the solution of actual, relevant educational problems than would have been possible as isolated individual researchers. The programmatic approach creates a body of specialists, each concerned with the total evaluation model but by training and background interested also in specific elements of a total model.

Center researchers see themselves as individuals in a community of scholars who are interested in the study of evaluation, and whose individual knowledge and competencies will be applied to problem solution as part of a team. The activities are diverse, yet inter-locking; the respective roles of each researcher become more effective in the field of evaluation within the programmatic structure and approach.

PEOPLE

These, then, are the principle goals of CSE and its strengthened programmatic thrust for achieving the goals. But lofty goals and organizational procedures are not enough. The major resource of CSE is its people.

While all Center personnel are important, and play a special role in forming the Center, CSE is especially proud of its research staff. A major part of what CSE is centers around the intellectual colleagueship of a number of professors and researchers, all productive scholars, who consider the Center as the locus of their research activities. The members of this group have selected the Center, and have been selected by the Center, because they have faith in its purposes and its ability to achieve these purposes.

TERMINATING PROJECTS

TERMINATING PROJECTS:
A FOREWORD

CSE began operation in the summer of 1966 dedicated to the study and improvement of evaluation. Its first research efforts were organized around the analysis of evaluation which followed directly from its model of evaluation, which had three major aspects:

1. *Instructional variables;*
2. *Contextual variables;*
3. *Criteria of instruction.*

The first two years of CSE operation were devoted to the study of the variables in each of these three crucially important areas in evaluation studies. This two year phase was necessary because educational research was not mature enough to indicate the measures and variables vitally important in evaluation of education.

It is now appropriate for the Center to use the results of this initial research in the development of evaluation models and systems, and to terminate the first stage in the development of CSE's research. As a result, the first stage programs on the following pages are being terminated and their results applied to the development of the new programmatic approach introduced elsewhere in this report.

INSTRUCTIONAL VARIABLES PROGRAM

THE INSTRUCTIONAL VARIABLES PROGRAM

(U.S.O.E. Projects 0501 through 0505)

Co-directors: M. C. Wittrock and M. H. Jones

TO EVALUATE THE LEARNING, retention and transfer resulting from the instruction of elementary school children, this CSE program studied two major areas: (1) individual differences in children's memory and perceptual traits as these affect classroom learning; (2) the effect of the organization of the instructional stimuli on the results of instruction.

A study during 1966-67 evaluated the effects of teaching programmed reading in an elementary school by using measures of learning, retention and transfer. Next year the research group of this project will progress from the experimental study of evaluating instructional variables to the evaluation of day-to-day classroom instruction in naturalistic settings. The instructional variables program has served its purpose of helping the Center learn important instructional and human variables to measure in ongoing classroom teaching.

SHORT TERM MEMORY FOR ITEM AND ORDER

(U.S.O.E. Project 0501)

THIS PROJECT WAS DESIGNED to assess the importance of individual differences in several memory abilities among children, as they affect classroom instruction.

THE FIRST PHASE of the project, the studies of visual-verbal and visual-figural memory among second and fourth grade children, were completed, comparing verbal with figural memory, and memory for ordering of items with memory for items alone. The data are now being analyzed.

ABSTRACTNESS OF MATERIALS IN LEARNING

(U.S O.E. Project 0502)

PROJECT 0502 EXPLORED the structure of materials in classroom learning. Specifically, the program was aimed at discovering whether there is a relationship among units, classes and systems of knowledge and, if so, whether this system can be used as a given hierarchy for the purpose of evaluating materials for their ease in learning and retention. These data are also being analyzed.

GENERALITY OF RULES AND THEIR SEQUENCING IN SOLVING PROBLEMS

(U.S.O.E. Project 0503)

FOLLOWING UP EARLIER WORK on organization of instructional variables in evaluation projects, the effect of generality of rules for transfer was studied. This was an effort to determine what kind of rules yield transfer. The purpose was to discover whether general rules or specific rules are superior to each other or to examples in promoting transfer of learning. These data are also being analyzed.

FORMATION OF PERCEPTUAL UNITS

(U.S.O.E. Project 0504)

THE PURPOSE OF THIS PROJECT was to measure individual differences in the ways in which, and the extent to which, elementary school children can be taught to group perceptual units. Groups of fifth graders were trained by three methods and the efficacy of these training procedures evaluated. The data of this project are now being statistically analyzed.

DEVELOPMENT OF HIERARCHICAL ORGANIZATION IN MEMORY STORAGE

(U.S.O.E. Project 0505)

TO DECIDE WHETHER it is important to index the degree of hierarchical organization of instructional materials in evaluation studies, the effects of organization upon learning and remembering was investigated in this project.

FOUR GROUPS OF FIRST and third graders were given lists of words to recall which could be organized in varying degrees. There was a very significant increase in total recall with age, as expected. The older children did not organize memory any better than the younger ones on the first trial, but improved significantly over five trials, whereas the first graders did not improve. Clustered lists, where the words in the same category occurred in immediate sequence rather than scattered at random, were the most effectively organized lists. In evaluating school instruction, indices should be obtained to measure the hierarchical organization of instruction.

CONTEXTUAL VARIABLES PROGRAM

CHILDREN'S OBSERVATIONAL LEARNING IN THE CLASSROOM

(U.S.O.E. Project 0506)

Project Director: Seymore Simon

THE MAJOR PURPOSE OF THIS STUDY, in terms of the role and effects of classroom recitation, was to compare the direct learning of original behavior in the natural classroom environment with learning from the experience of other children.

THE APPROACH WAS TO USE two school districts--seventh and eighth grade students being drawn from the first district, third and fourth graders from the second. A total of 16 different classes were employed. Intact classes were used, and the normal classroom teacher acted as the experimenter. Two classes within each school district served in any one condition, thus minimizing chances of group treatment confoundings.

Three experimental groups (a direct group, a vicarious group, and a direct and vicarious group) and a control group were used. After all groups received pretest measures consisting of word association, unusual uses and theme writing, pupils in the direct group were asked to give repeatedly different associations to new lists of words over three consecutive days of training. Pupils in vicarious groups listened to the uncommon associations of their classmates over the same period of time. The direct/vicarious groups received both direct and vicarious experience over the three days; the control groups received neither.

Transfer was assessed in terms of response to new word stimuli over the consecutive days, as well as to new unusual uses, items and a new theme on the final day. The word association and unusual uses responses were scored according to previously established norms. The themes were scored for type-token ratios and rated for originality. To establish reliability of ratings for theme orginality, two raters worked independently and achieved an interrater reliability of 95 percent agreement.

THE DATA HAVE BEEN PREPARED and are now in the process of analysis. A report will be completed by December, 1968.

IMITATION OF AND PREFERENCE FOR TEACHERS VARYING IN REINFORCEMENT STYLE, AS A FUNCTION OF AGE, SOCIAL CLASS AND PERSONALITY

(U.S.O.E. Project 0601)

Project Director: Norma D. Feshbach

A MAJOR OBJECTIVE OF THIS PROJECT was to extend to a broader population the findings of an earlier study in which learning traits of disordered boys from advantaged and disadvantaged backgrounds were the subjects. The findings of that study indicated that boys from advantaged backgrounds showed significantly more imitation of a teacher than comparable boys from disadvantaged backgrounds. In addition, incidental imitation of a teacher who used positive reinforcement was significantly greater than incidental imitation of a teacher who used negative reinforcement. The findings suggest both a process and an outcome factor, i. e., imitation and resulting social behaviors. These have important implications for the evaluation of classroom programs.

The present project is intended to partially replicate the earlier study, as well as studying the imitative influences on both girls and boys in a more typical population of school children from middle and lower socio-economic backgrounds.

THE METHOD EMPLOYED was to use as subjects children in one advantaged and one disadvantaged school. The sample included 120 boys and girls, 60 Negro and 60 Caucasian.

THE RESULTS PROVIDED STRIKING CONFIRMATION of the principal experimental hypotheses. Imitation of the positively reinforcing teacher was significantly greater than of the negatively reinforcing teacher. Advantaged Causcasian children exhibited reliably more imitative behavior than disadvantaged Negro children. Within each social class, girls imitated more than boys. In addition, imitation and dependency were positively correlated in the sample of advantaged children.

Papers resulting from this project included a presented paper at the American Educational Research Association, February, 1968; a detailed abstract of the paper published in the Association's Paper Abstracts for 1968; and an unpublished master's thesis at UCLA.

LEARNING STYLES I: CONSISTENCIES AND RESPONSIVENESS TO POSITIVE AND NEGATIVE REINFORCEMENT AS A FUNCTION OF SOCIAL CLASS AND PERSONALITY

(U.S.O.E. Project 0602)

Project Director: Norma D. Feshbach

THE PURPOSE OF THIS STUDY was to ascertain whether children show a consistent mode of response over a variety of tasks to specific types of reinforcing procedures. A second goal was to determine personality and social class correlates of any consistencies reflected in children's behavior.

THE METHOD EMPLOYED WAS TO PRESENT CHILDREN with various sequences of four reinforcements over two sessions.

THE MAJOR FINDING of the study was that, for the advantaged group, the effects of positive and negative reinforcement on learning is dependent on the sequences of reinforcement.

Middle class children who initially received praise in the sequence of four reinforcements (sequence: praise, criticism, criticism, praise) did significantly better under punitive conditions than under praise.

Those middle class children who initially received criticism in the sequence (sequence: criticism, praise, praise, criticism) did significantly better under praise than under criticism. This interaction was most marked for the most difficult learning task.

Though lower class children tended to perform somewhat better under conditions of praise under both sequences, the differences are not statistically significant.

This project was completed in April, 1968. A master's thesis resulted and a complete report is currently being prepared.

TEACHING STYLES IN FOUR-YEAR OLDS

(U.S.O.E. Project 0603)

Project Director: Norma D. Feshbach

THE PRIMARY PURPOSE of this project was to determine whether the variations in learning (both direct and indirect) observed in other responses to different reinforcement treatments are reflected in the style of "teaching" which children use in instructing other children.

INCLUDED AS SUBJECTS in this study were 204 three and four year old children, drawn from nine different nursery and Head Start programs. The boys and girls who participated in the study represented four different social class-ethnic groupings: Advantaged White Children, Disadvantaged White Children, Advantaged Negroes, and Disadvantaged Negroes.

After the four year olds had been taught how to solve a puzzle by the experimenter, they in turn were required to teach three-year-olds how to solve the puzzle. In every case, the social class and race of the four-year-old "teacher-child" was identical to that of the three year old "pupil". However, the sex of the two was systematically varied. The primary dependent variable was the number of positive and negative reinforcements emitted by the four year old "teacher" to his three year old "pupil".

THE RESULTS INDICATE that advantaged white children use positive reinforcement significantly more often when instructing their "pupils" than do their counterparts in the other three groups. Within group comparisons indicate that middle class Negro children, lower class white children and lower class Negro children all use critical and negative remarks significantly more often than they do positive reinforcement and praise. The differences for the middle class white children are not significant and in the opposite direction.

A paper on this study was presented at the American Educational Research Association meeting in February, 1968 and will be reprinted in Readings in the Social Psychology of Education, Allyn and Bacon, 1969. In addition, a paper is in press to be published in Child Development in 1969.

THE INFLUENCE OF TEACHER'S PERSONALITY AND VALUES UPON THE EVALUATION OF CHILDREN VARYING IN PERSONALITY AND CLASSROOM BEHAVIOR

(U.S.O.E. Project 0604)

Project Director: Norma D. Feshbach

THE FOCUS OF THIS INVESTIGATION is on the evaluation of the relationship between personality characteristics and the values of teachers relative to their preferences for and expectancies of differing types of children.

THE STUDY INCLUDED 237 female student teachers as subjects. All were engaged in elementary school student teaching. Three measures were administered to each subject: The Edwards Personal Preference Schedule, a Semantic Differential for assessing teacher values and attitudes, and a Classroom Situation instrument for assessing teacher evaluations of children displaying a variety of behaviors.

The situation test consisted of 16 classroom situations depicting children displaying one of four personality constellations--for example, non-conforming, dominant, flexible as one constellation; passive, orderly, dependent as another. There were two situations for boys and two for girls representing each of the four personality groups, yielding a total of 16 situations. Using six point rating scales, the teachers evaluated the child depicted in each situation, and also indicated the degree to which they would prefer to have the child in their class.

FINDINGS THUS FAR HAVE RESULTED from analyses of the Classroom Situation instrument and the Edwards Personal Preference Schedule.

The empirical findings on the situation test showed that student teachers rate the dependent-submissive child significantly more favorable than the challenging-independent child. Further, a factor analysis of the instrument revealed one factor with two poles to account for the four personality constellations. In addition, the instrument appears to measure a single underlying personality trait.

The findings on the EPPS suggest that this group of student teachers respond on the instrument very similarly to the norm group. This part of the study is currently being completed and put into final form.

FUTURE PLANS FOR THE PROJECT include completion of the analyses of past work, and an extension of the project to investigate the relationship between student teachers' attitudes toward students and the manner in which such attitudes can be modified.

Specifically, the study will attempt to determine whether student teachers given a favorable first impression and subsequent conflicting negative information about a child will take a more negative attitude more rapidly than will subjects given an initial negative impression followed by subsequent conflicting positive information. Two measures will be used. One will consist of two situations describing an elementary school age child engaging in favorable and unfavorable behaviors. The second measure will describe a child with positive or negative academic records. Ratings will be made on five behavioral dimensions. In addition, the changed data will be related to personality variables.

The entire study, including the extension, will be completed by January, 1969.

THREE PAPERS ON THE PROJECT have been written to date, one published in _Educational and Psychological Measurement_, 1969; a second is in press for the _Journal of Educational Psychology_, 1969; and a third has been submitted for presentation at the American Educational Research Association meeting in February, 1969.

TEACHING STYLES OF MOTHERS, AND TEACHING AND LEARNING STYLES OF THEIR FOUR YEAR OLD OFFSPRING

(U.S.O.E. Projects 0605 & 0606)

Project Director: Norma D. Feshbach

THREE PRIMARY PURPOSES MOTIVATED these combined studies. The first is replication of Project 0603, Teaching Styles in Four Year Olds, which indicated that advantaged white children's use of reinforcement in instructing younger peers differs significantly from that of advantaged Negro children and disadvantaged Negro and white children. The second purpose was to relate the style of reinforcement of four-year-olds from various socio-ethnic backgrounds to their mothers' use of reinforcement. The third purpose was to relate the children's use of reinforcement to their responsiveness, to differential incentive conditions.

The studies in this project are concerned with an important individual difference variable relevant to evaluation--namely, the use and responsiveness to differential reinforcement.

FOUR-YEAR-OLDS DRAWN from four different socio-ethnic backgrounds instructed three-year old "pupils" on a puzzle-solving task. Subsequently, these four-year-olds were taught a different puzzle by their mothers. The third task involved a concept formation problem, in which half of the four year olds learned the problem under positive incentive conditions, and the remainder learned the problem under more negative conditions.

ALTHOUGH THE DATA ANALYSIS is not yet complete, preliminary findings indicate that the ethnic background of the mother and child are significantly correlated with the use of negative reinforcement. The data analysis will be completed by December, 1968, and a complete report and journal article will be prepared.

IMITATION OF TEACHER PREFERENCES IN A FIELD SETTING

(U.S.O.E. Project 0607)

Project Director: Norma D. Feshbach

THE PRINCIPAL OBJECTIVE OF THIS PROJECT is to extend the laboratory studies of incidental imitation to an actual classroom situation. More specifically, the study is aimed at evaluating indirect influences of teachers upon children's attitudes and preferences, and ascertaining individual context variables mediating these effects. In terms of its broader implications, the project extends the range of behaviors that should be assessed in the evaluation of educational programs.

THE EXPERIMENT WAS CARRIED OUT in two classrooms. A control sample, drawn from several other classrooms, was also included. Teachers' incidental comments, evaluating photographs of animals, were systematically varied and the effects upon pupil preferences were subsequently ascertained.

DATA DESCRIBING INDIVIDUAL DIFFERENCES in response to teacher influence are currently being analyzed, and a paper is in preparation. The expected completion date is December, 1968.

ACHIEVEMENT MOTIVATION AND THE ATTRIBUTION OF RESPONSIBILITY

(U.S.O.E. Project 0610)

Project Director: Bernard Weiner

THIS STUDY WAS DESIGNED to investigate the relationship between need for achievement and the perceived locus of responsibility for success and failure. The underlying theory guiding this experiment is that the ascription of causality precedes and mediates achievement motivation. The dependent variables are intended to assess whether a child primarily perceives the achievement outcome as caused by himself (internal locus of control), or by some aspect of the environment (external control).

THE SUBJECTS IN THIS STUDY were children in the second to sixth grades. They were engaged in tasks varying in difficulty at which they either succeeded or failed, depending on the experimental condition. In addition, the experiment includes various indices of achievement motivation to aid in the development of a valid test of achievement concerns for children.

A FINAL WRITTEN REPORT is currently being prepared.

THE EFFECTS OF INTERRUPTION VERSUS FAILURE ON SUBSEQUENT PERFORMANCE

(U.S.O.E. Project 0611)

Project Director: Bernard Weiner

THE PURPOSE OF THIS PROJECT was to separate the effects of task failure from simple task interruption on subsequent task performance.

THE PROJECT METHOD included classifying college students according to strength or resultant achievement motivation (need for achievement, minus test anxiety). All subjects fail to complete an achievement-oriented activity. Under one condition, incompletion is due to a personal failure. In the other condition, the incompletion is attributable to an environmental event.

Intensity of performance and choice between various achievement-related activities were measured subsequent to the personal or environmental block. Differential reactions between the achievement-oriented groups were expected, as a function of the source of frustration.

A REPORT OF THE PROJECT is currently being prepared.

COUNSELING: AN INSTRUCTIONAL MODEL

(U.S.O.E. Project 0612)

Project Director: Garth Sorenson

THE GOAL OF THIS PROJECT was the development of a means, a conceptual framework, and instruments and procedures for evaluating the outcomes of school counseling. Counseling is viewed as a special instructional process designed to bring about behavioral changes in students relating to two traditional educational goals, i.e. developing both freedom and responsibility in students.

THREE INSTRUMENTS TO EVALUATE counseling have been developed and partially tested as part of the study. One is a measure of psychological stress, for use with student teachers. The second is a measure of counselee reactions to the interviewer. The third is a measure of self-concept for use with high school students. The self-concept measure is being used in several projects involving minority group students in the Los Angeles city schools.

A PRELIMINARY CONCEPTUAL FRAMEWORK was described in a CSE Occasional Report of 1967. Another report, now being published, describes the first application of the framework. It details some of the problems arising when a rigorous attempt is made to evaluate counseling in a natural situation, rather than in a laboratory setting, and presents some solutions to these problems.

A report is being prepared, and a paper will be submitted to a professional journal. The project will be completed in December, 1968.

INSTRUMENTATION FILE AND MANUAL

(U.S.O.E. Project 0613)

Project Director: Norma D. Feshbach

TWO PRINCIPAL TASKS ARE INVOLVED in this project. The first was to develop a central file of relevant individual context variables, including measuring instruments and literature references. The second was the compilation of a manual.

The following individual context variables have been included in the first version of the manual. These variables were selected on the basis of their theoretical and empirical relevance to the study of the evaluation of instructional programs. The variables are:

- affiliation
- aggression
- anxiety
- attitudes toward school
- cheating
- conformity
- constricted vs. flexible control
- curiosity
- delay of gratification
- dependency
- dogmatism
- dominance submission
- empathy
- expectancy (locus of control)

field articulation
imitation
impulsivity vs. reflection
intolerance of ambiguity
introversion-extroversion
leveling sharpening
masculinity-femininity
moral judgement
need achievement
popularity
risk taking
self concept
similarity-dissimilarity
social desirability
suggestibility

The manual is organized under relevant variable headings. Within each variable a definition, recommended measures, descriptions of the measures, measure evaluation and references are included.

THE MANUAL HAS BEEN COMPLETED. However, periodic revisions will be made as new data become available. A report on this project has been written and is currently in press.

EVALUATION FRAMEWORK FOR THE EDUCATIONAL BEHAVIOR OF MINORITY PUPILS

(U.S.O.E. Project 0614)

Project Director: C. Wayne Gordon

TO CREATE A FRAMEWORK for the evaluation of educational behavior of minority pupils, several studies of factors related to the aspirations and achievement of minority pupils have been completed.

FACTORS EXAMINED INCLUDE pupil characteristics, social background, family influence, values, and school context. Sixty-four variables were studied, providing some understanding of those variables most strongly related to educational outcome of minority pupils.

A SUMMARY PAPER will be prepared setting forth the findings from the previous analyses which have special relevance to the evaluation of educational programs.

It is clear that the strength of the relationships of selected factors to achievement and aspirations varies with ethnicity. For example, school context indexed by ethnic density is more strongly related to the reading comprehension of Mexican-American than of Anglo pupils. Also, parent aspirations for educational attainment are related to educational achievement for both Anglos and Mexican-Americans, but not equally so.

The summary of these findings will suggest a minimum set of variables which are related to the achievement of each of these separate groups: Anglo, Mexican-American, Negro, and Oriental.

DIFFERENTIAL EFFECTS OF TEACHER CLASSROOM MODES BY PUPIL CHARACTERISTICS

(U.S.O.E. Project 0615)

Project Director: C. Wayne Gordon

THIS PROJECT EXPLORES the effects of the teacher leadership mode on students from varying socio-economic, racial, and ethnic backgrounds.

TEACHER LEADERSHIP MODES ARE BEING MEASURED by the instrument developed by Gordon, Adler, and McNeil, in their study, "Dimensions of Teacher Leadership in Classroom Social Systems." Teacher leadership modes are derived from teacher scores on three leadership dimensions--task orientation, authority orientation and expressive orientation--and put into a three-dimensional matrix.

The student variables are peer group structures, which are measured sociometrically, and attitudes. Four attitude scales are used to assess self-esteem, school orientation, independence from peers, and orientation toward parental authority. Multivariate analyses are being used, as are controls for socio-economic status, race and ethnicity.

DATA WILL BE COLLECTED within the next few weeks and analysis will be finished by December, 1968.

RELATIONS BETWEEN COMMUNITY CONTEXT AND THE ORGANIZATIONAL CLIMATE OF THE TWENTY LEAGUE SCHOOLS

(U.S.O.E. Project 0616)

Project Director: C. Wayne Gordon

THIS EXPLORATORY STUDY OF THE RELATIONSHIPS between the community context of a school and the school's own organizational climate used the League of Cooperating Schools as a sample.

COMMUNITY CONTEXT MEASURES were collected from the U. S. Bureau of the Census. School attendance boundaries were matched with census tracts to attain measures of community context. The community variables were: 1) socioeconomic status; 2) urbanization; 3) segregation; 4) age; 5) stability. Organizational climate was measured by Halpin and Croft's *Organizational Climate Description Questionnaire*. Cross tabs of the community variables with the subtest scores of the *OCDQ* were analyzed. A re-analysis of the data with school district size controlled was also carried out.

THE MAJOR FINDING OF THE STUDY is that principal behaviors are influenced by community context to a greater extent than teacher behaviors.

The project will soon appear as a CSE publication, which is now in the final stage of preparation.

CONTEXTS OF STUDENT ALIENATION

(U.S.O.E. Project 0617)

Project Directors: C. Wayne Gordon and Carl Weinberg

THE MULTIFOLD PURPOSES OF THIS STUDY link student expectations and perception of the environment to alienation within the school. One aim of the study was to conceptualize student alienation in terms of various areas and behaviors which reflect student integration into the activities of the school. Another purpose was the development of an instrument from which student perceptions of the value of specific activities and the behaviors of teachers in the classroom can be derived. This instrument also gives some idea of the student expectations of teachers, and of student concepts of the standard teacher role in their school.

THE INSTRUMENT TO MEASURE these factors has been organized from sample items which were previously generated. A rationale for the development of a classification scheme to describe student alienation has been completed. Alienation was used basically as an output measure for schools and systems, and was also linked to several context variables, as well as to achievement.

Student role expectations also were translated into a set of items which have been organized into an instrument.

FINAL QUESTIONNAIRES for Roles and Student Alienation were completed as of September 15, 1968.

CRITERION MEASURES PROGRAM

STUDY OF THE CHANGES IN ACHIEVEMENT AT THE LEVEL OF THE INDIVIDUAL ITEM

(U.S.O.E. Project 0703)

Project Director: Rodney W. Skager

THE AIM OF THE LAMMP Diagnostic Achievement Test, developed for use in the Los Angeles Model Mathematics Program, was to diagnose group performances on a variety of learning objectives in mathematics. Such information was expected to be more useful in evaluating the effects of an instructional program than are the total scores on the usual standardized tests, particularly where the latter are based on items unrelated to the instructional goals of the program being evaluated.

The research group also hoped that experience in developing such measures in the evaluation of a curriculum development program would provide at least a tentative model for the construction, use and analysis of similar instruments in other evaluation contexts.

THE LAMMP DIAGNOSTIC TEST WAS ADMINISTERED, along with a number of other measures, to about 700 seventh grade students in three junior high schools.

The students, of normal aptitude, but at least one year behind in mathematics achievement, were participants in a curriculum development program. Approximately half of the students were assigned to experimental classes utilizing the new curriculum, with the remainder attending mathematics classes using the regular curriculum.

WHILE TOTAL SCORES of the LAMMP Diagnostic Achievement Test and the arithmetic sub-test of the Iowa Test of Basic Skills failed to reveal differences between experimental and comparison groups, analysis of subscores derived from Diagnostic Test items of similar content and/or requiring a similar process in solution did reveal different patterns of achievement gain between experimental and comparison groups. The observed differences were in part validated by classroom observations of instructional content and by teacher ratings of item relevancy.

This project showed that achievement tests can reveal differences in learning results when (a) items are selected on the basis of specified instructional goals, (b) data is analyzed at the level of homogeneous item groups, and (c) other information about the instructional process is used to corroborate the test findings.

Experience gained in this research was a primary source for the conception of the Instructional Objectives Measurement System.

COMPLETION DATE of the project is December 31, 1968.

DEVELOPMENT AND USE OF MEASURES OF LEVEL OF COGNITIVE DEVELOPMENT IN EVALUATION OF INSTRUCTIONAL PROGRAMS

(U.S.O.E. Project 0704)

Project Director: Rodney W. Skager

INITIAL WORK ON THIS ASPECT of the Los Angeles Model Mathematics Program (LAMMP) surveyed the development of materials and procedures for measuring cognitive skills and the use of these measures as part of the evaluation of a curriculum development in mathematics. It was reasoned that cognitive skills such as conceptions of quantity, grouping strategies, seriation and cardination operations, and formal reasoning abilities should be closely related to the ability to profit by instruction in abstract subject matter.

If these relationships are found to exist, conclusions can be drawn as to the relative effectiveness of LAMMP instruction for students at different levels of cognitive development. Evidence for such interactions between student characteristics and instructional conditions is potentially of great significance for evaluation methodology. Instructional methods do not ordinarily work equally well for all types of students. Specifying the groups for which a program is or is not effective would be a valuable contribution of any evaluation research.

THE RESEARCH INVOLVED the individual administration of eleven measures of cognitive skills to approximately 100 seventh grade students of normal aptitude who were at least one year behind grade expectation in mathematics achievement. A variety of other measures were administered to these students, including a standardized achievement test, a diagnostic test in mathematics constructed for the project, and a questionnaire on demographic and related socio-economic information. The two achievement measures were administered at the beginning and end of the school year.

PRELIMINARY ANALYSES OF THE DATA have revealed positive correlations between total scores in mathematics tests and several of the cognitive measures. The latter was mainly confined to measures of formal reasoning. Quantity conceptions as elicited by tests of conservation did not appear to be related to achievement, although in view of their age level, a surprising proportion of the subjects were unable to pass these measures.

Data analyses now under way are searching for relationships between gains in specific types of mathematics performance and particular cognitive skills. The research group will also inspect relationships among the cognitive measures in the hope of arriving at a smaller set of relatively independent tests.

REPORT ON THE EFFECTIVENESS OF THE LAMMP PROJECT FROM THE VIEWPOINT OF INDIVIDUAL CRITERION MEASURES

(U.S.O.E. Project 0705)

Project Director: Rodney W. Skager

THIS PROJECT WAS AIMED at evaluating the Los Angeles Model Mathematics Program (LAMMP) during its first year. During this period, the LAMMP staff developed and used a mathematics course utilizing instructional materials and procedures appropriate for junior high school students of normal aptitude who are at least one year behind grade expectation in mathematics. LAMMP features programmed materials, feedback systems, and a mathematics laboratory in a program directed toward the needs of the inner-city student--frequently a member of a minority group.

THIS EVALUATION RESEARCH DESIGN required dividing some 700 students in the seventh grade of three junior high schools between experimental classes using the new materials and comparison classes keeping to regular procedures and materials. Students were administered a standardized achievement test, a diagnostic mathematics test developed especially for the program, a questionnaire on social and demographic factors, and a test of attitudes toward school--especially toward the study of mathematics.

A smaller sample of 100 students was individually examined on a variety of cognitive skills. Both achievement measures were given at the beginning and end of the school year.

THE PRELIMINARY REPORT OF FINDINGS, issued in August, 1968, emphasized that teachers of the experimental classes had geared the development of their materials to strengths rather than weaknesses in mathematical achievement, as revealed by pretest data on diagnostic test items. Taking into account the entry skills of eligible students should have been the first step taken in the development of the LAMMP curriculum.

The same findings applied in a lesser degree to the comparison classes. There was no evidence of overall differences between experimental and comparison groups with respect to gains in mathematics achievement between pre and post testing. Different patterns of gains were evident for certain subgroups of items, suggesting that experimental and comparison programs in part emphasized different content.

The final report on the project, incorporating additional information (see projects 0703 and 0704), is due in January, 1969.

PHYSIOLOGICAL CORRELATES OF PUPIL ACHIEVEMENT AND ADJUSTMENT

(U.S.O.E. Project 0706)

Project Director: William H. Lucio

THIS STUDY TESTS THE HYPOTHESIS which holds that measures of individual differences in the functioning of the autonomic nervous system (ANS), whether the differences are inherited or acquired, provide a basis for the prediction (and evaluation) of pupil achievement and adjustment. The study also related data from standardized tests of individual physiological functions of the ANS to indices of pupil academic achievement (including "over-and under-achievement"), adjustment, and physical and mental health.

METHODS USED TO GATHER DATA for this study involved obtaining physiological measures on all available subjects enrolled in the University Elementary School at UCLA. A Behavior Inventory, adapted from the Fels Behavior Inventory and developed for this study, has been administered to 255 pupils with two to four teacher raters per subject. In addition, verbal and performance measures (Bijou Wide Range Achievement Test, Stanford Achievement Test, S/scores yielding I. Q. performance ratios, and ratings on an Achievement Index) have been obtained on all subjects.

RESULTS FOR COMPUTER TREATMENT were available at the end of August, 1968, as data reduction of all physiological variables was completed at that time. Expected date of completion for the study is December, 1968, and a report will be ready in early 1969.

THE EFFECTIVENESS OF DIFFERENT INSTRUCTIONAL PROCEDURES USED IN THE TEACHING OF YOUNG CHILDREN

(U.S.O.E. Project 0707)

Project Director: Carolyn Stern

THE PROJECT IS DESIGNED TO DEVELOP INSTRUMENTS to evaluate the effectiveness of different instructional procedures used in the teaching of young children.

THE INVESTIGATION STUDIES THE RESPONSE made by children of 4-8 years of age, who represent a variety of socio-economic levels. The children respond to two types of instructional material which have been administered. One, a series of pictures, includes realistic as well as abstract and non-objective forms. The other, a set of nonsense disyllables, samples across phonemes in initial, medial, and terminal position.

A card for each response has been prepared from which linguistic coding was completed for each response word. The code for each response for each child is being assigned.

THE STUDY IS TO BE COMPLETED by December, 1968.

DEVELOPMENT OF INSTRUMENT FOR ASSESSING EGO STRENGTH

(U.S.O.E. Project 0708)

Project Director: Louise L. Tyler

THE AIM OF THE PROJECT was to develop an instrument for assessing ego strength. It was felt that criterion measures are an aspect of evaluation that need investigation, and that ego strength is a significant variable. Psychoanalytic psychology was chosen as the framework for instrument development for two reasons: (1) Ego strength as a concept has been most thoroughly studied by the psychoanalysts. (2) Their conceptions of ego strength seem the most salutary.

THE INSTRUMENT WAS DEVELOPED FOR USE at the elementary school level. It took the form of 12 cards presenting stick figures, in which the situation represented aspects of the school environment, such as teachers, children, and subject matter. A group of 28 children at the University Elementary School took the test.

The design of the scoring key was based upon the idea that formal characteristics are more important than interpretations. Scoring was done by the director and a research assistant.

THE FOLLOWING TENDENCIES were noted among the children of ages 4-6, 6-8 and 10-12. The youngest seem to have the most difficulty in satisfying desires; the kind of elaborations indicated seems to become more various with age --possibly due to a buildup in defenses, and so the complexity of ego. There seem to be more opposite responses in the youngest and oldest--perhaps related to concern and insecurity about autonomy at these ages.

This pilot study has been more concerned with the validity than the reliability of the instrument. Both, however, need much more investigation. An annotated bibliography for the project is available.

SCHOOL BOARD EVALUATIONS OF INSTRUCTIONAL PROGRAMS

(U.S.O.E. Project 0709)

Project Director: Jay Scribner

THE BASIC OBJECTIVE OF THIS PROGRAM has been the consideration of standard measures suitable for use in evaluating the instructional programs of both individual schools and school systems.

ATTENTION WAS DIRECTED toward the affective domain of the evaluation process, as influenced only by the school-system standards of achievement. Since factors that influence the goals of a specific school system are numerous and complex, an effort must be made to determine these goals--or at least to establish a sound basis on which to classify system goals.

DATA ARE BEING ANALYZED and a final report on the follow-up study is being rewritten. Completion is set for December, 1968.

REVIEW OF ORGANIZATIONAL CRITERION MEASURES

(U.S.O.E. Project 0710)

Project Director: Marvin C. Alkin

THIS STUDY WAS A REVIEW TO DETERMINE organizational criterion measures which might prove to be potentially applicable for use in evaluating programs.

SELECTION AND IDENTIFICATION of additional criterion variables for another study were to be based on these preliminary findings. These variables were to be tentatively tested in a field-study program involving high schools recently accredited by the Western Association of Schools and Colleges. At the termination of this activity, a decision was reached as to whether further investigation would be worthwhile.

THE PROJECT IS NOW TERMINATED.

BENEFIT-COST STUDIES AND PROGRAM BUDGETING IN PUBLIC SCHOOLS

(U.S.O.E. Project 0303)

Project Director: Erick L. Lindman

COST-EFFECTIVENESS and program budgeting have come under study by the Center during its first two years of operation. This effort has been concerned primarily with (1) the identification of programs and subprograms for budgeting purposes and (2) development of an expenditure classification structure which will provide information for program cost analyses.

A NATIONAL CONFERENCE on program accounting was held at UCLA on July 15, 16 and 17, 1968, and attended by representatives from major organizations and school systems interested in the subject, including the U. S. Office of Education.

One of the purposes of the conference was to analyze problems and issues encountered by U.S.O.E. in preparing a new accounting manual for public schools. "This kind of endeavor", wrote Glenn Boerrigter, Chief of the U.S.O.E. Organization and Administrative Studies Branch, "will make a tremendous long-range impact on the planning program and budgeting for our elementary and secondary schools".

Allan Lichtenberger, primarily responsible for producing the new accounting manual, called the conference "most effective.... Among the most fortunate disclosures", he wrote, "was the extent to which attention has been turned to accounting by program. It is the major reason for the revision of Handbook II. I have a fairly comfortable feeling that the suggested structure of accounts is in accord with the primary concern, that of program accounting."

WITH THE PUBLICATION of the proceedings of this conference, the Center is terminating this project. U.S.O.E. has recently made two major grants for the purpose of developing program budgeting procedures. Continued effort in this program budgeting area would probably duplicate these new efforts. Several other center projects will continue to emphasize concerns for cost-effectiveness evaluation and management information systems. (See projects 0203 and 0307).

NEW AND
CONTINUING PROJECTS

NEW AND CONTINUING PROJECTS:
A FOREWORD

On the basis of the knowledge and experience derived from its earlier programs, CSE has been able to develop a new, more cohesive and comprehensive programmatic approach to achieving its goals.

All new and continuing research projects at the Center fall within one or another of the three new programs: Evaluation of Instructional Programs, Evaluation of Educational Systems, and Evaluation Theory and Methodology.

The Program Chart on page 67 shows the specific projects in each of these program areas.

Following is a brief description of the aims and activities of the current CSE programs:

PROGRAM ON THE EVALUATION OF INSTRUCTIONAL PROGRAMS

PURPOSE

The goal of this program and the projects within it is to construct a model and methodology for evaluating instructional practices, programs and procedures.

Instructional programs that would involve such evaluations include educational innovations and developmental activities such as curricula, methods of instruction, technological aids, and day-to-day classroom teaching. This program differs from educational research in general since its focus is on: 1) evaluation of formal instructional programs and classrooms (as opposed to general behavioral science research on learning, personality, etc.); and 2) on providing a means of gathering information about instruction for educational decisions.

RESEARCH ACTIVITIES

The first step in evaluating teaching and instruction involves obtaining measures in three categories:

1. Instructional materials and the characteristics of the teaching;
2. Salient individual differences of the learners;
3. Results or criteria of instruction, such as achievement, transfer, retention, and student interest.

These three types of measures enable one to take the *second* step in evaluation: relating the results of instruction of different types of students to the instruction and teaching they received in school. Unlike common studies of evaluation of instruction which measure only the results of instruction, the CSE procedure is designed to evaluate the instruction *per se* --the essential evaluation problem classroom teachers and school principals face every day.

The Center's procedure does not stop at measuring student learning or achievement. Instead, it is designed to measure how classroom instruction and teaching have contributed, or failed to contribute, to student learning with different groups of students of varying backgrounds and intellectual processes.

PROGRAM ON THE EVALUATION OF EDUCATIONAL SYSTEMS

PURPOSE

This program is directed toward construction of an evaluation model and methodology for studying educational systems, such as schools or school systems. Thus, the focus is on ways of studying the total organization or system, rather than on a particular program.

RESEARCH ACTIVITIES

The program will include the evaluation of inclusive components of systems or total systems, such as schools or school districts.

Here, the focus usually does not involve a specific set of procedures, but is related to how adequately the total organization or system is functioning. This type of evaluation research is less frequently represented in general educational research, though in method and conception, it is analagous to comparative studies utilizing statistical, rather than experimental, control.

Contemporary examples from the evaluation field include the National Assessment Project, the International Study of Mathematics, and the Coleman Report.

PROGRAM ON EVALUATION THEORY AND METHODOLOGY

PURPOSE

Construction of a general evaluation model for identifying and integrating evaluation problems that should be studied is the aim of this program.

The model's functions would include generating criteria for determining the relevance of other programs to the task of building a model of evaluation. This would also include construction of submodels as needed. In addition, it is necessary to invent and put into operation new data gathering methods and data analysis models.

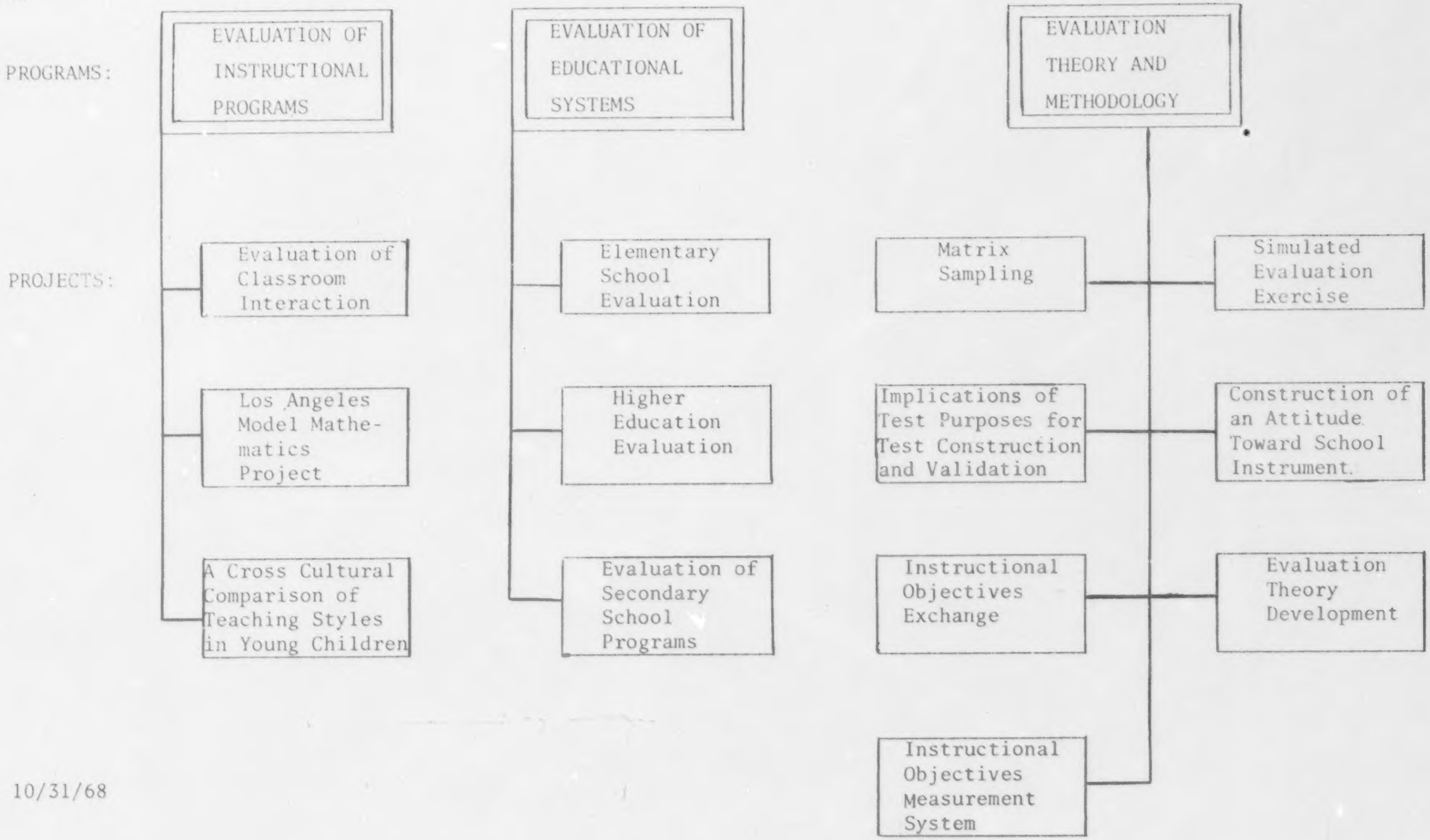
RESEARCH ACTIVITIES

Projects in this program might include reviews of the literature of evaluation studies, determining what changes are likely to occur in education that would affect an evaluation model, and expanding this document.

Additionally, study projects that deal with improving the technology of evaluation might be conduted. Examples of this work include matrix sampling and behavioral objectives.

CENTER FOR THE STUDY OF EVALUATION

PROGRAMS AND PROJECTS


PROGRAMS:
EVALUATION OF INSTRUCTIONAL PROGRAMS
EVALUATION OF EDUCATIONAL SYSTEMS
EVALUATION THEORY AND METHODOLOGY
PROJECTS:
Evaluation of Classroom Interaction
Los Angeles Model Mathematics Project
A Cross Cultural Comparison of Teaching Styles in Young Children
Elementary School Evaluation
Higher Education Evaluation
Evaluation of Secondary School Programs
Matrix Sampling
Simulated Evaluation Exercise
Implications of Test Purposes for Test Construction and Validation
Construction of an Attitude Toward School Instrument
Instructional Objectives Exchange
Evaluation Theory Development
Instructional Objectives Measurement System

10/31/68

PROGRAM ON THE EVALUATION OF INSTRUCTIONAL PROGRAMS

LOS ANGELES MODEL MATHEMATICS PROJECT

(U.S.O.E. Project 0301)

Project Director: C. Wayne Gordon

CREATING NEW MATHEMATICS INSTRUCTIONAL MATERIALS and developing an instructional program to meet the needs of disadvantaged average seventh grade math pupils are the dual objectives of this project.

To accompany these general objectives, special tests were developed for pre and post test assessment of pupils' ability in mathematical conceptualization and in the curriculum. In relation to these tests, charge-sensitive test instruments to observe pupil change will be developed. To add information about the pupils, data on an extensive number of social items was collected by observation techniques which yielded pupil's aspirations, values, attitudes, and achievements.

In addition to test development, studies were made of the interaction of individual pupil characteristics with the instructional aspects of the program, of the functions of non-standard language among minority pupils related to instruction, learning and evaluation, and of the administrative context and its effects on the instructional and evaluation program.

TEACHERS IN THREE Los Angeles City junior high schools located in disadvantaged areas volunteered to cooperate in this experimental mathematics program.

Beginning in September, 1967, evaluation teams conducted pretesting in six to eight classes in each school, and in approximately the same number of comparison classes. Post-testing with the same groups was done in April and May, 1968. Several test instruments were used: the Iowa Test of Basic Skills, the Los Angeles Model Mathematics Project Diagnostic Test (developed from the instructional goals of the project), a number of measures of the level of cognitive development, a social survey instrument, and an attitude toward school test.

A PRELIMINARY EVALUATION REPORT, submitted to the Los Angeles City Schools in August, 1968, presented the analysis of data obtained from the various test instruments used. The report pointed out a limitation faced by the report team in that reports from school personnel implied students had not been randomly assigned to experimental and control classes. In addition, data from all three tests indicated that the groups were not approximately equivalent at the outset.

The most significant findings emerged from an examination of the relationships between the proportion of students passing each item on the Diagnostic Test given in the fall and the ratings made by teachers at the end of the year on the relevancy of each item to their instructional program. Teachers had evidently been directing instruction to those areas in which students were initially more capable.

Another publication generated by the project, Field Studies Series, No. 1, the CSE Simulated Evaluation Exercise, presents a design to train evaluators working the field settings. Two masters' theses were also developed from this study.

IMMEDIATE PLANS ENTAIL THE COMPLETION of the final report by December, 1968. The report will include results of analysis of items on the Diagnostic Test, analysis of the measures of cognitive level and attitudes toward school, and an additional analysis of the social survey items.

Continued analysis of data obtained from field study, but outside the objectives of the report to the Los Angeles School System, is anticipated.

A CROSS CULTURAL COMPARISON OF TEACHING STYLES IN YOUNG CHILDREN

(U.S.O.E. Project 0608)

Project Director: Norma D. Feshbach

THE MAJOR PURPOSE of this study is to extend cross culturally the research relating to teaching styles of children and their mothers (Projects 0603, 0605, 0606) and to determine the generality of the psychological concomitants of sociological variables revealed in the earlier studies.

THESE INITIAL STUDIES indicated that the children's and parents' use of reinforcement (an individual difference variable very relevant to classroom learning) is related to the socio-eco-ethnic background of the individual. It was found that four year old advantaged white children use positive reinforcement significantly more often when instructing three year old children in a puzzle task than do disadvantaged white children or advantaged and disadvantaged Negro children.

Within group comparisons also indicated that middle class Negro children, lower class white children, and lower class Negro children all use critical and negative remarks significantly more often than positive reinforcement and praise, while the differences for the middle class white children are insignificant and in the opposite direction. The second two studies in this series were concerned with another important individual difference variable relevant to evaluation, namely, the use and responsiveness to differential reinforcement.

In these studies four year old children drawn from four different socio-ethnic backgrounds instructed three year olds on a puzzle solving task. Subsequently, these four year old boys and girls were taught a different puzzle by their mothers. The third task involved a concept formation problem. Half the four year olds learned the third task under positive incentive conditions, with the remainder learning the task under more negative conditions.

The results indicate that ethnic backgrounds of the mother and child are significantly correlated with the use of negative reinforcement.

THE CURRENT PROJECT will involve middle and working class British and Israeli four year old boys and girls and their mothers as subjects. Arrangements have been made in both countries. Data will be collected in these two countries between October, 1968, and February, 1969. Reports and other publications will be prepared by June, 1969.

THE EVALUATION OF CLASSROOM INTERACTION

(U.S.O.E. project number assigned)

Project Co-Directors: M. S. Wittrock &
M. H. Jones

THIS RECENTLY INITIATED PROJECT is directed toward providing information about an important aspect of instruction which is not now represented in our evaluation system, and is not normally present in evaluation studies of elementary schools.

THE SIGNIFICANCE OF THE PROGRAM is based on the theoretical reasons it offers for asserting the importance of interaction occurring in the classroom. These theories may be summarized as:

1) Transmission of information over any communication channel always results in some loss of information. The extent of the loss in the communication of information from teacher to pupil has not been measured with any accuracy, nor have the conditions determining the loss of information been stated.

2) Events occurring during a learning process ordinarily interact to such a significant

degree that learning outcomes cannot be reliably predicted from events taken singly. The intermediate step between the initial measurement of such variables as individual abilities of the students, characteristics of curricula and teaching materials, sociocultural background of students, characteristics of the school physical plant and the administrative climate, and the final measurement of outcomes of education is therefore an important process in itself, and will contribute additional information to the evaluation system from direct measures of the interaction.

3) The Center's commitment to a theory of evaluation which is multivariate and interdisciplinary, and which seeks for explanation of the findings of evaluation studies, must build into its system for evaluation some means of tracking down these explanations.

There are several empirical reasons for seeking some of the explanations in the classroom instructional process:

1) The Center's study on the educational achievements and aspirations of Mexican-American youth found that the school context was a significant contributor to success, along with socio-economic status and

ethnic group, but could point to no specific factors which might be held accountable.

2) The many excellent evaluations of initial reading programs have found that the "teacher-variable" overrides the differences in instructional materials ("teacher-variable" referring vaguely to something occurring in the classroom instruction).

The importance of introducing instruments for evaluating the classroom interaction is clear. Although some aspects of this interaction have been studied, they are largely confined to measurement of social and emotional interactions, and scales are available for these purposes. The neglected aspect of the interaction lies in the areas of information transmission, communication of meaning, and influence on the thinking processes.

THE APPROACH OF THE PROGRAM is to operate in a real evaluation setting. The work will be done in two school situations which are part of larger evaluation studies so that relevant data concerning other predictive factors which may operate in conjunction with the classroom measures and measures of the outcomes of the total educational event are available, and can be viewed together. Two immediate tasks will be initiated:

1) The development of an existing research instrument on the form of classroom interaction to a usable index for the Elementary School Evaluation Program--one which can utilize short samples of classroom behavior, can be readily coded, and is reliable--and development of a cost-effective instrument for the evaluation of ongoing classroom instruction.

2) The development of a cost-effective instrument to measure congruence of concepts between teacher and pupils, as a means for determining community of meaning associated with the language used in the classroom. This instrument will be built upon existing research into concept meaning, associations, and concept hierarchies. The multi-station Student Monitoring System will be used to collect data on both the sequence of associations the student makes to the teacher's concepts and the speed with which he does so. Ultimately, paper-and-pencil tests of concept congruence will be prepared as part of the kit of evaluation materials for the Elementary School Evaluation Program.

THE PRODUCTION SCHEDULE for the development of these instruments is as follows:

1) The index of the form of classroom interaction will be ready for use by the Elementary

School Evaluation Program by December 31, 1968; revision will follow the second total evaluation test. The final analysis of instrument effectiveness is to be completed by December 31, 1969. The final report will be ready by March 31, 1969.

2) The instrument to measure concept congruence will be designed by January 30, 1969, the test materials ready for use in test runs by March 1, the first information from actual classroom evaluation by April 1. Revisions, subsequent tests, and evaluation of the instrument as part of a total evaluation package will require the summer and early fall, to be completed by October 1, 1969. The instrument will be administered to a larger number of classrooms in the fall, 1969, and analyzed in conjunction with other variables in the evaluation study. Analysis of the data and the final report will be ready by March 31, 1969.

THESE INSTRUMENTS ARE INTENDED for two purposes:

(1) For use by administrators or evaluators in evaluating schools and systems in such a way as to indicate not only over-all quality, but also possible trouble spots;

(2) For use, perhaps in a different form, by teachers themselves to evaluate their own performance and gain insight into where the communication problems lie with a particular class.

PROGRAM ON THE EVALUATION OF EDUCATIONAL SYSTEMS

ELEMENTARY SCHOOL EVALUATION STUDY

(U.S.O.E. Projects 0201, 0202, 0203)

Project Director: Stephen Klein

THE TWO MAJOR OBJECTIVES OF THIS PROGRAM are (1) to develop a system for obtaining accurate, sensitive, reliable and salient information about an elementary school that will facilitate decisions regarding a school's program, policies and plans; and (2) to construct a system that will be cost effective and permit comparisons between schools on several relevant dimensions.

It is also expected that the process of working towards these two objectives will clarify fundamental issues and concepts in evaluation theory. This may occur through the identification of critical variables and their cause and effect interrelationships.

THE MULTITUDE OF POTENTIAL VARIABLES have been categorized into four classes: (1) instructional (such as nature of content and degree of individualized instruction); (2) organizational (such as the formal and informal relationship between teacher, student and principal; (3) contextual (such as school facilities and the student's home environment); (4) criterion (such as student achievements and attitudes).

A number of measures within each of these classes have been obtained. This was accomplished by searching the literature for relevant measures, revising existing instruments, and in some instances, developing new tests and questionnaires. Although most of these measures are administered to students, some are designed for teachers, principals and parents. Certain statistical and methodological procedures for integrating and analyzing the information obtained from these measures have also been developed.

A PRELIMINARY TRYOUT of some of the new measures was conducted in December, 1967. This study, involving fifth grade students in two elementary schools, helped to improve many of the questionnaires and procedures. A more extensive tryout was conducted in May, 1968. It involved about 1,000 students (grades four to six) in five elementary schools. The data obtained in this later study is expected to be analyzed by February, 1969.

THE PLANS FOR NEXT YEAR center about the construction of an "evaluation kit". The purpose of this kit is to provide a preliminary tool for implementing the information system. Designed for use by school personnel, the kit will contain a set of measures selected on the basis of their relevancy to salient evaluation issues. The kit will also contain the procedures for giving the measures, collecting other relevant information such as grades, scoring and preparing data for analysis procedures, and guides to interpreting the information for educational decisions. The kit will be used in the sample evaluation of some 15 schools before being put into large-scale operation, and refined and enlarged in successive years with the information obtained.

EVALUATION OF SECONDARY SCHOOL PROGRAMS

(U.S.O.E. Project 0302)

Project Director: Marvin C. Alkin

THE OBJECTIVES OF THIS PROJECT are: (a) development of a mathematical model which simulates the characteristics of the secondary school; (b) construction of a series of mathematical models explaining the effects of the school's characteristics upon its performance outputs; and (c) development and/or employment of various ancillary techniques, such as cost-effectiveness, to aid in the interpretation of the explanatory models.

IN APPROACHING THE STUDY, raw data supplied by the Western Association of Schools and Colleges for 100 California four-year high schools was coded, keypunched and "cleaned", and financial, ethnic and geographic data added. From the data, 103 study variables were generated. The study then proceeded through the following steps:

The first phase of analysis (primarily concerned with the examination of "orienting" techniques) obtained a statistical description of the total sample in terms of the study variables. This listing provided an "overview" of how the sample behaved on each of the study characteristics, and indicated variables which were too rare to be trustworthy.

The second procedure was a "contingency analysis" of the performance of the schools on the study criteria. This procedure was designed to indicate the basic relationships of each explanatory variable with each school performance criterion.

The exact nature of these basic relationships was probed by the third procedure. Linear, quadratic and cubic equations descriptive of each relationship were developed by computers, their accuracy compared, and the most nearly correct selected for further study. Many of the relationships were found to be non-linear and will require a modification of the linear multiple regression model to be used in later analysis.

The final procedure was a by-product of the third, in that it used graphs of relationships produced at the same time as the equations. These graphs described the basic single-variable relationships with the performance criteria and seemed rather superior to the "contingency" analyses produced in the second procedure.

THE DATA BASE for phase two of the processing was expanded by the addition of data from 303 more four-year high schools obtained by the WASC, and data obtained from student questionnaires. The total data now involves 404 high schools and approximately 140 study variables. The data has been coded and keypunched, and is currently being "cleaned".

FUTURE PLANS are to have the latest data processed and subjected to the procedures of phase one analysis in the early months of next year.

The first explanatory model to be examined for each of the criterion measures will be the linear multiple regression model, modified by the introduction of quadratic and cubic terms of the explanatory variables.

Additional multivariate analyses, such as canonical correlation, will then be evaluated as to their ability to improve upon, or augment, the regression model. This will be followed by one or more specific applications of the basic explanatory mode.

High priority will be given to performing a "cost-effectiveness" analysis of the secondary school in a bid to determine the optimum allocation of resources for a given school situation. Depending upon the findings at tnat time, the project group anticipates a decision on the future of the program by September, 1969.

EVALUATION IN HIGHER EDUCATION

(U.S.O.E. Projects 0401, 0402, 0403)

Project Director: Dr. C. Robert Pace

BALANCE AND BETTER INFORMED JUDGEMENTS about higher education in the United States depend upon the range and relevance of criterion and contextual measures that are used in an evaluation program. The application of a broad range of such measures in a national field study will provide the data for studying better evaluation techniques, and arriving at more balanced conclusions.

THIS NATIONAL APPRAISAL OF HIGHER EDUCATION is primarily concerned with a range of outcomes of higher education that may be observable in the behavior of college educated adults and students, together with the nature of the educational experience, the characteristics of the institutions, and with various background and personal characteristics of the individuals. In order to accomplish this appraisal, questionnaires have been devised for administrators to sample alumni, upperclassmen, and freshmen at approximately 100 colleges and universities across the country.

Criterion measures included in the questionnaires cover a broad range of involvement of the individual in contemporary society and culture, measures of knowledge about and attitudes towards certain major changes taking place in American society, personal judgements about the extent to which higher education contributed to various outcomes, and other values attached to the college experience. The questionnaire also includes a variety of personal background information.

Contextual measures of institutional characteristics will be applied which are pervasive, stylistic, and felt throughout a single campus, and which also differentiate one campus environment from another. Three measures have been completed: measures of campus morale, of quality of teaching and faculty-student relationships (both operational), and a third variable, an analysis of student logs to see how students spend their academic time. These will be translated into a brief questionnaire.

PLANS CALL FOR a testing schedule to begin in January, 1969, with a mailing of the questionnaire to alumni, followed by the administration of the questionnaire to juniors during February and March and to freshman in September.

Beginning work has been done on the development of at least one model test, which will measure the extent students can combine their knowledge about various disciplines to thinking about complex social issues: human ecology, urban renewal, war, peace, and others.

A measure of elementary school environments similar to the College and University Environment Scales has been completed. It is hoped to expand this to include high school environments so that it will be possible to plot the changing environment in which education occurs and study the implications for longitudinal evaluations.

As a result of CSE activities in higher education, three doctoral dissertations have been awarded. Other publications include: The Measurement of Campus and Student Morale, by Lora Robinson and Richard Seligman; Evaluation Perspectives, 1968, by C. Robert Pace; Faculty And Curriculum As Measures of College Environment, by James M. Richards, Jr., and Richard Seligman.

IN THE NEXT YEAR, MANAGEMENT LOGISTICS and data processing will be a major focus of the project, with continuing work on new contextual and additional criterion variables.

Additional contextual measures to be developed are as follows:

1) Analysis of five different colleges, their organizational functions (labelled work, boundary stability, adaptive, and control).

2) Development of a measure of institutional stance toward student discipline and freedom.

3) Measurement of student morale, similar to those devised by the Armed Forces and industry.

4) Development of measures of institutional anonymity.

5) Measures of campus peer group patterns and influences.

The possibility exists that a liaison will be contracted to improve institutional research with the California State College system on a consulting basis. In addition, exploratory work was done on the content and design of item sampling methods for periodic opinion polling of college students which can be done during the time of student registration.

The League for Innovation in the Community Junior College, under the direction of Dr. Lamar Johnson, has requested that project to serve as consultants for evaluation of such innovations as they undertake.

Under consideration, but still in the exploratory state, is the possibility of some kind of official connection with one or more international evaluation programs. The relevance of such enterprises for the Center was investigated by Dr. Pace while on sabbatical in Europe.

PROGRAM ON
EVALUATION THEORY AND METHODOLOGY

A SIMULATED EVALUATION EXERCISE

(U.S.O.E. Project 0304)

Project Director: Marvin C. Alkin

TO INCREASE THE FLEXIBILITY OF EVALUATORS as they respond to constraints encountered in actual field conditions is the general goal of this simulation exercise. To achieve this goal, the exercise will provide an opportunity for teams of participants acting as evaluators to construct and subsequently learn to modify evaluation designs.

Through the simulation exercise, an evaluator will discover the necessity of being flexible in his evaluation design because controls designed into the original model may either break down or perhaps never even be implemented when used in actual conditions. Participants will experience the real problems and frustrations that occur in a field study.

IN THE EXERCISE, AN EVALUATION TEAM has contracted to evaluate three specially funded junior high school mathematics program over a period of one year. These are experimental mathematics programs introduced to a group of underachieving students. Participants in the exercise play the role of Center evaluators and are asked to carry out evaluation of this simulation exercise. The evaluation team is expected to submit the best possible final evaluation design, considering the constraints of actual field conditions throughout the simulated school year. All information appearing on the exercise is based on real data gathered from schools in which an experimental mathematics project actually was introduced. At the close of the exercise the participants are given the opportunity to discuss and evaluate their experiences.

THE PROJECT PLANS are to create from the exercise, when completed, a package including five major parts: 1) An annotated bibliography; 2) An instructor guide and material supplement; 3) A mock final report; 4) A simple version of the exercise; 5) A "how-to-do-it kit".

The project is currently in the process of compiling an annotated bibliography which will be sent to participants prior to their becoming involved in the exercise. This bibliography will include a summary of the current theories, methods, and techniques pertinent to the evaluation of educational programs. The instructor's guide and material supplement for the simulated exercise are in final stages of revision and should be ready for printing shortly. A mock final report, modeled after one of the Center's current evaluation programs, is being developed for distribution to participants after their participation in the Simulated Exercise. This report will provide detailed information concerning various activities, methodologies, and techniques that UCLA actually employed in the conduct of an educational evaluation. The report will also serve as a model for individuals who have not presented the results of an educational evaluation to a sponsoring agency. The report itself will not be a part of the exercise, but is designed rather as a by-product, so that upon leaving those involved will have had the experience of participating in a simulated evaluation in addition to acquiring the annotated bibliography.

The package of bibliographies, manuals, and mock report will represent one form of a training unit which might be employed to increase the efficiency of potential evaluators. However, several additional elements being developed will allow for flexibility of unit construction for application to different aspects of educational evaluation.

INCREASED FLEXIBILITY WILL BE GAINED by providing a short form or simple evaluation exercise. This exercise will essentially be developed as the longer form was. That is, an evaluation problem will be selected which will require a methodolcgical design capable of handling the objectives of evaluation. This simple exercise will be designed to allow participants to "quickly" participate in an evaluation endeavor without using sophisticated methodologies or tackling complex and difficult problems. Once completed, this simple form can be employed as the initial point in a week-long training session on evaluation which would be supplemented by speakers and lectures and terminated with the longer exercise.

The simple "How-to-do-it-Kit", to be developed as another element of the total training system, will identify important factors to be considered before undertaking an evaluation. It would list necessary prerequisites for conducting an evaluation study, provide sample evaluation designs which would fit various evaluation situations, and provide a step-by-step procedure outline for conducting an evaluation.

Potential trainers of evaluators will have at their disposal a basic evaluation unit composed of the annotated bibliography, long form of a simulated evaluation exercise, and a mock final report. In addition to this basic unit, they will have available a short form of the simulated exercise and a "How-to-do-it Evaluation Kit". These elements can be combined at the discretion of the trainers to create packages capable of meeting their training needs. Factors such as time, size of group, and experience of participants will determine how these additional units will best serve the learning situation.

A FINAL PHASE OF THIS PROJECT is the development of an assortment of evaluation exercises and games appropriate for use by small groups or individuals. They are designed specifically to provide understanding and insights into other aspects of evaluation problems, and to be readily available and utilizable with a minimum of preparation.

CONSTRUCTION OF AN ATTITUDE TOWARD SCHOOL INSTRUMENT

(U.S.O.E. Project 0609)

Project Director: Norma D. Feshbach

ATTEMPTS HAVE BEEN MADE BY SOCIOLOGIST and educators to measure and assess the attitudes which children and adolescents hold toward various aspects of school life. None has proven to be completely adequate. It has, however, been empirically demonstrated that attitudes toward school are a significant variable relating to school performance and adjustment. Thus, a clear need exists to develop a valid and reliable measure to assess Attitudes Toward School.

Further, when young children or culturally disadvantaged groups are the target population, verbal measures are inappropriate. Therefore, the purpose of this study is to construct a non-verbal Attitude Toward School instrument which will appeal to pupils ranging from 10-15 years of age.

INSTRUMENT CONSTRUCTION INVOLVES a lengthy and rigorous procedure of initial item selection, analysis, revision, and empirical testing until a final, useful, reliable, and valid measure for general use has been developed.

The measure has been rationally constructed in the following manner: Original items were selected from a larger pool of items intended to cover three general areas of school activities. These are: 1) general school including classroom, playground, assembly, and teacher-child interactions; 2) math activity, including scenes relating to math content areas; and 3) social studies, including scenes relating to social-studies content areas.

Stimulus materials consist of 30 pen and ink drawings, each of which depicts items related to one of the three areas described above. The cards are individually administered to the subjects and presented one at a time. The subject indicates on a seven point scale how much he would like to do what the student in the picture is doing. A board containing the seven point scale is placed in front of the subject, who points out the degree of liking by selecting one of the seven points. Responses are then recorded by the experimenter. The seven point scale ranges from extremely positive (+++) to extremely negative (---).

TO DATE, THE MEASURE has been administered to LAMMP subjects and a middle class comparison group. The data are currently being analyzed.

The Attitude Toward School instrument is still in a preliminary stage and it is expected to take another year before instrument construction is completed. It is expected that revisions in item selection and item content will be made and the measure administered to additional samples of subjects. It is also anticipated that ultimately a measure, useful with junior high school age boys and girls of all ethnic and socioeconomic groups, will result. The entire project is to be completed by December, 1969.

In addition to developing the Attitude Toward School instrument itself, a report discussing the project will be written.

MATRIX SAMPLING IN EDUCATIONAL RESEARCH

(U.S.O.E. Project 0701)

Project Director: Theodore R. Husek

WORK IN MATRIX SAMPLING, often referred to as item sampling, has been theoretical during this year. As a result of the empirical studies completed the year before, it was necessary to change the approach to correct problems uncovered in past studies. These empirical studies showed the matrix sampling variance estimate was negative under certain conditions. Stated another way, negative internal consistency among items was a result under certain conditions.

DEALING WITH THE ACTUAL DERIVATIONS and assumptions underlying the matrix sampling formulas became a necessity. The derivations were extremely complex and tedious, the assumptions not clearly stated. This led researchers to consider other possible mathematical formulations of the technique. This was necessary to facilitate the derivations of primary formulas, explicate the assumptions underlying the derivations, including implications for both binary and non-binary items, and draw other implications with regard to the use of matrix sampling on various kinds of tests, e.g., norm referenced vs. criterion-referenced achievement tests.

IT WAS FOUND that casting matrix sampling into an analysis of variance framework made it possible to accomplish three objectives. A paper new in the final stages of preparation will attempt to demonstrate the following points:

1) Matrix sampling can be viewed as a simple two factor, random model analysis of variance design, the matrix sampling formulas for estimating the mean and variance being simply the analysis of variance formulas for estimating components of the underlying linear and additive model.

2) In contrast to non-binary items, binary items can be expected to violate homogeneity of variance and independence assumptions. When the between-subjects variance is less than the residual variance, a negative matrix sampling variance estimate results.

3) Given the use of binary data, the usual achievement test items, in contrast to objective-meeting or criterion-oriented test items, may be more amenable to matrix sampling methodology with respect to variancc

estimation. (Estimation of the mean appears to be unaffected by the type of items or test used).

4) Of the various attempts to deal with negative variance component estimates in multiple matrix sampling, either the maximum liklihood method of equating these estimates to zero or the method of Winsorizing the distribution of estimates appear to be most promising. A simulation study is necessary to determine the shape of the distribution of variance component estimates for matrix sampling, as well as the relative efficiency of the two methods for handling negative estimates.

AT THE PRESENT TIME A MANUAL on how to use matrix sampling is being written. This manual will incorporate the project's work to date, as well as the endeavors of other researchers. It will hopefully be sufficiently explicit so that the relatively unsophisticated evaluator can use matrix sampling procedures in a sensible fashion.

IT IS PROPOSED IN THE NEXT YEAR to investigate at least two additional problems. The possible presence of context effects in the sampling of test items will be studied. It is expected this evaluation program will require an experimental study to begin early in 1969.

Investigation into the optimal examinee-item sample size ratios depending upon statistical characteristics of the examinee-item population will be the second problem. This project will require a simulation study. At the present time work is being done on a computor program to generate random data matrices according to pre-specified characteristics.

DIFFERENT PURPOSES FOR TESTS AND THEIR IMPLICATIONS FOR TEST CONSTRUCTION AND VALIDATION

(U.S.O.E. Project 0702)

Project Director: Theodore R. Husek

THE PURPOSE FOR WHICH a test is constructed has important implications for how it should be constructed and judged. This is particularly important as the idea of criterion referenced tests receives increasing attention from evaluators.

In the past, the procedures of classical test theory have traditionally been applied to norm-referenced tests. Thus, tests were designed to maximize the variability among those tested. However, criterion-referenced tests are designed to minimize inter-subject variance and concentrate on learned behavior of those tested.

The identification of problem areas related to criterion-referenced test construction has been the focus of the study this year.

DIFFERENTIATING BETWEEN IMPLICATIONS of purposes of tests, it was found, can be implemented by analysis of the following problem areas, using, as an example, norm-referenced and criterion referenced tests.

1. Variability. Since the meaningfulness of a norm-referenced score is basically dependent on the relative position of the score in comparison with other scores, the more variability in the scores the better. With criterion-referenced tests, variability is irrelevant. The meaning of the score is not dependent on comparison with other scores; it flows directly from the connection between the items and the criterion.

2. Item Construction. When the test constructor writes items for norm-referenced tests, he is interested in achieving a balance between item difficulty and item discrimination to maximize variability among subjects. The criterion-referenced item writer is guided by a wholly different goals: the item must represent the class of behaviors delimited by the content objectives.

3. Reliability. The internal consistency of a norm-referenced test can be easily measured by some statistic like coefficient alpha. But these statistics rest on the assumption of the presence of individual differences, i.e., variability. A criterion-referenced test could be quite reliable (all

items are measuring the attainment of the same objective); yet such indices dependent on variability might not reflect this internal consistency.

4. Validity. Again, studying the validity of a test is mainly a process of correlational analyses, and these analyses are also dependent upon the assumption of variability. Thus the notion of content validity becomes particularly important in the context of criterion-referenced measurement.

5. Interpretation of test data. With respect to norm-referenced tests, the methods of interpreting the results of an individual's performance are well known. Group-relative descriptors such as percentile rankings or standard scores are commonly used. When interpreting an individual's performance on a criterion-referenced test, however, such group-relative indices are not appropriate. It would appear that scores which are essentially "on-off" in nature are more suitable. That is, it may be sufficient to report whether or not the learner has displayed a desired criterion behavior. Depending on how the test results are to be used, this criterion behavior may be further subdivided so that more than dichotomous decisions can made.

THESE IMPORTANT PROBLEM AREAS--test variability, item construction, reliability, validity, and interpretation of the test data--were the focus of a paper which discusses these areas in detail for criterion-referenced tests.

IN THE NEXT YEAR OF THE PROJECT, development or modifications of old psychometric procedures for handling the development, plus analysis of criterion-referenced test, will be accomplished. Other kinds of tests as defined by their purposes will also be treated.

INSTRUCTIONAL OBJECTIVES EXCHANGE

(U.S.O.E. project number unassigned)

Project Director: W. James Popham

THE INSTRUCTIONAL OBJECTIVES EXCHANGE has been established to serve as a national depository for criterion-referenced instructional objectives and related measurement devices, as well as a dissemination service to keep educators informed of work in the area of criterion-referenced instruction.

The primary service of the Objectives Exchange will be to make available alternative criterion-referenced objectives and measuring devices to assist school personnel in their instructional and evaluation activities. This service is a matter of volitional selection of instructional material by those users who participate in the Objectives Exchange program.

TO PERFORM THESE FUNCTIONS, a visible clearinghouse will be maintained to keep educators aware of the diverse instructional objectives development projects throughout the nation.

A bank-like service will be provided, whereby a school district or comparable educational agency can "draw out" all relevant measures for as many subjects, grades, and topics as desired.

The Objectives Exchange will also continually update, refine, and expand the pool of objectives and related evaluation measures for each field.

THE OPERATIONAL PLAN for conducting this program has been initiated. Formulation of the operational framework and drafting of the necessary literature to publicize the beginning activities of the program has been accomplished. This has involved the planning for collection, storage, processing, and distribution of objectives and evaluation measures. This system will be evaluated on the basis of input, processing problems, and related operational adjustment, and revised as required. At present it is planned to process, catalogue and commit to microfiche objectives and related test items as they are collected.

To some extent the users should undoubtedly help bear the cost of the objectives and evaluation measurements they receive. However, this is a complex practical problem, and any user fee plans will be worked out slowly and carefully.

MOST INSTRUCTIONAL EXPERTS would agree that while there are great dividends to be secured from having educators employ items which measure their success in achieving explicit objectives, this certainly does not assure that successful instruction will emerge. More often than not, just the opposite may be revealed. Yet, the detection of inadequate instruction represents a great stride forward.

To what extent the Objectives Exchange should follow up on the instructional tactics of its users or work with other agencies who are advocating such tactics has not yet been resolved.

INSTRUCTIONAL OBJECTIVES MEASUREMENT SYSTEM

(U.S.O.E. project number unassigned)

Project Directors: W. James Popham &
Rodney W. Skager

THE INSTRUCTIONAL OBJECTIVES MEASUREMENT SYSTEM will develop an objective-item measurement system for several subject fields. The selection of these subject fields will probably be in those areas recognized as critical, such as communication and quantitative skills.

Research personnel have been retained to develop criterion-referenced objectives and their evaluation measures. For each selected field the objectives will be specified, and related evaluation measures (not only test items but also diverse types of learner performance measures) will be developed.

Once a pool of objectives and test items has been developed and evaluated they will be made available to school personnel for use in instructional programs. These people will be the users of the system.

USERS OF THE SYSTEM will be required to select objectives, rather than individual test items. To avoid objectives that are too general, or items that could be listed under a broad objective and, therefore, create a test too long for normal classroom use, the system will be concerned with the specificity of items that it selects. In this regard, the system will collect user preferences for the various levels of specificity and user ratings on the relevancy of items under each condition of specificity. The latter will be utilized to ascertain whether the ratings of item relevancy are lower under more general sets of objectives.

A coding system will be developed for the retrieval of test items. This item code will include the general objectives under which the item is classified, the specific objectives which the items meet, the type item, and the number of the item.

OF PRIME IMPORTANCE IS feedback to the user and the kind of report generated by the system. As a standard practice, a pre and post test approach may be developed since a major advantage of the system will be its capability of informing teachers as to the entry skills of their students. Questions will arise as to whether providing users with initial item difficulties is enough, or whether it is possible to use the same items in both tests. Further, the system will be concerned with how feedback information can best be organized. These questions can be answered by a survey of preferences of users and observation of the instructional process under different feedback conditions.

THE ADVANTAGE OF THE SYSTEM is the service and information it supplies to the users. Teachers and evaluation researchers will be able to select behavioral objectives from material provided by the distributing agency. Further, they will receive appropriate testing instruments for measuring student performance on the selected objectives, and be informed on student progress in meeting these objectives.

EVALUATION THEORY DEVELOPMENT

(U.S.O.E. project number unassigned)

Project Director: Rodney W. Skager

INCLUDED WITHIN THIS NEW PROJECT are a variety of activities taking place within the Center related to development of evaluation theory. The project is in some ways more an administrative device for systematically considering several broad range interdisciplinary activities in the Center than an actual project.

THE SPECIFIC ACTIVITIES included as a part of this project involve the continuing work on conceptualizing evaluation on the part of a large number of Center researchers in a seminar-workshop devoted to this objective.

Several conferences similar to the conference held at the Center in December 1967, but narrower in scope, will be held as part of this project. Research assistants assigned to the project are providing much of the substantive input to the delineation, discussion, and theory development by assisting in bibliographic research of related activities.

ANOTHER MAJOR ACTIVITY to take place within this project is related to monitoring the work being done on management information systems as an evaluation procedure, and providing a mechanism for the sharing of information with those also concerned about this vital area.

It is not appropriate to provide precise time schedules for a project such as this, which has an its concern synthesis, theory building, and conceptualization.

ADMINISTRATION

REORGANIZATION AND REDIRECTION

A major reorganization of CSE management has come into effect in recent weeks. The purpose: to make administration compatible with program structure.

NEW DIRECTOR, NEW ADVISORS

In September, the appointment of Marvin C. Alkin to the Center Directorship was announced by John I. Goodlad, Dean of the UCLA Graduate School of Education. Merlin C. Wittrock, was named Chairman of the Center's new Program Planning Board. In that role, he will carry the title of Co-Director. Dr. Alkin had been an Associate Director since the Center's inception in June, 1966, and was a member of the group that prepared the funding proposal for the Center.

The administrative alignment at CSE gives the Director a tightly coordinated team designed to both recommend and implement program policies. As chief administrative officer, the Director maintains full responsibility for program decisions; but crucial assistance in making these decisions is now provided by the seven-member Program Planning Board.

In a recent letter to Board members, Dean Goodlad defined the Board's functions in this way:

> In his role as Chairman, Merl Wittrock is responsible for guiding the inquiry and discussion which will lead to the initiation of new projects and the termination of old ones. Constantly, you'll need to be striving for unifying concepts which provide inclusion/exclusion criteria for what the Center is all about. The Program Planning Board will formulate and transmit recommendations to Marv Alkin, who will serve the Board as executive officer somewhat in the way that a superintendent of schools serves the board of education.
>
> But there is one critical difference. The Board is advisory to Marvin, but he must assume responsibility for the decisions of the Center....

Members of the Program Planning Board are appointed by the Dean of the School of Education. It is CSE's intention that the Board represent a relatively broad range of points of view with respect to the activities of the Center, and that it represent the active working members of the organization (though not necessarily excluding other individuals). It is conceivable that a portion of the Board will be elected by the research staff members of the Center in future years.

The current membership of the Board is as follows: Merlin C. Wittrock, Chairman; Dean C. Wayne Gordon, Erick L. Lindman, C. Robert Pace, W. James Popham, Rodney W. Skager and Louise L. Tyler.

ADMINISTRATIVE ORGANIZATION

There are two other major changes in the administrative organization of the Center. First, the appointment of a full time Assistant to the Director is anticipated. Second, in conformance with the reorganization of the Center's research activities into three major programs, CSE anticipates placing more responsibility with the three program directors, who will be identified later.

The Director is a faculty member of the Graduate School of Education and is on a part time appointment with the Center. Since he feels the necessity for spending a major portion of his time in working with the research programs of the Center, there is a need for a full time Assistant to the Director to supervise many of the day-to-day administrative functions. The Assistant will have primary responsibility for managing a number of phases of the Center's operations. Reporting directly to him, administratively, will be a Senior Administrative Assistant in charge of fiscal management, university procedural problems, and accounting records. The Coordinator of Statistical Services will be responsible to the Assistant. In addition, individuals responsible for dissemination services and clerical services will work most directly with this individual. It is expected that the Assistant will also assume much responsibility for liaison with related organizations, and will provide assistance to the Director in research project coordination.

With this assistance, the Director will be able to focus a major portion of his energies on the primary tasks of monitoring the research programs and providing substantial intellectual leadership in effecting appropriate changes in the Center's research activities.

In all its reorganization moves, CSE has sought to create a management system directly correlated to the nature of its work. It has concluded that the burdens of proper management require that the administrative functions of the Center be centralized, with the Director having clear authority to make decisions, and with an administrative support structure that allows him to delegate as much responsibility as possible. Moreover, CSE feels that policy decisions must be based on the broadest and most expert advice possible.

This solid organizational foundation--depicted on the chart at the end of this section--enables CSE to construct a future of increasingly productive work.

CSE RESEARCH STAFF

While there are a number of full time professional researchers in the Center, the more typical pattern is to involve a variety of faculty members who envisage the Center as the main locus of their research activities.

In terms of budget allocations, the typical faculty member is engaged in R & D Center research between 2/9 and 4/9 of his time during three quarters of the year, and in full time research during one quarter of the year. In actual practice, where the university teaching load is moderate (five courses in three quarters), most CSE researchers spend even larger portions of their time in research activities in the Center, including the direction of related doctoral dissertation research. Many key Center research personnel have reduced teaching loads, supported by CSE, and teach only one course per quarter.

CSE has recently added a number of highly qualified research personnel to its staff. Stephen Klein, formerly with the Educational Testing Service, Princeton, New Jersey, joined the Center July 1 as a full time research psychologist. CSE is indeed fortunate to have added Dr. Klein to its staff. During recent weeks, W. James Popham, a prominent member of the faculty of the Graduate School of Education, has been added to the CSE staff. His willingness to now join CSE represents a very major addition to the Center.

Clarence Bradford and James Trent are other faculty researchers recently added to the staff. Getting these men was dependent on a high level of cooperation from the Graduate School of Education, in providing related faculty positions for researchers needed for the CSE program. Dr. Bradford was previously at Washington University, St. Louis. His interests are in evaluation, statistics, and research methodologies in the behavioral sciences. He has most recently been working on canonical regression analysis techniques. Dr. Trent brings a social psychologist's perspective to the study of problems in the evaluation of higher education. He came to UCLA from the Berkeley R & D Center in Higher Education.

Several other new faculty appointments are becoming active participants in the Center. Eva Baker, whose specialty area is curriculum and evaluation of instruction, has worked in recent years at the Southwest Regional Laboratory. David O'Shea, an educational sociologist with interests in systems evaluation, came to UCLA from the University of Chicago.

All of these appointments, as well as several others presently pending, are designed to provide increased strength to a CSE staff that has already demonstrated high capability.

RELATIONS WITH OTHER AGENCIES

CSE maintains a close working relationship with the Research Division of the Institute for Development of Educational Activities (IDEA). The two groups are cooperating in the development of the evaluation system for elementary schools. After CSE has defined the variables, instruments, and preliminary procedures for collecting data, the plan calls for IDEA to collect the actual data from the member schools of the League of Cooperating Schools. These data will then be processed by the Center. It is also expected that IDEA will participate generally during the development of the evaluation system, and will contribute to the formulation of concepts and theory--as well as to the resolution of problems.

In addition, plans are being formulated to involve selected schools outside of California to insure that the developing evaluation system will be applicable and effective on a national basis.

During the past year, CSE has had contact with a number of agencies interested in Center research activities related to specific agency problems. In all instances, the Center has attempted to provide whatever advice was possible.

In some cases, discussions with various agencies about the possibility of performing certain functions are still in progress. For example, CSE has been approached by the Title III office (U.S.O.E.) with respect to the use of the simulated evaluation materials in training Title III Directors. The Center has hastened its schedule on the development of these materials in order to fulfill this request. The simulation materials might also be used in fulfilling a request for teaching evaluation techniques to heads of teacher training programs funded by the Educational Professions Development Act. Furthermore, several discussions have taken place between administrators of the Center and Dr. David S. Bushnell of U.S.O.E. about the feasibility of the Center participating in the evaluation of the Educational Systems for the Seventies program.

In addition, the Center has provided advice to various school districts and to the state of California program on desegregation.

While the Center has been quite willing to provide the advice and consultation which the agencies desire to improve their educational evaluations, it has been generally reluctant to engage in field evaluations of a strictly field service nature. The Center has been guided in the selection of its activities by a consideration of the importance of the study in terms of its potential contribution to a theory of evaluation, methodology and technology of evaluation, and national visibility and generalizability.

The Center has also, of course, maintained close relationships with a number of Regional Laboratories, and has been in contact with other U.S.O.E, R & D Centers, as well as with private research organizations whose interest areas overlap those of the Center.

SPACE

The R & D Center occupies an area in the Graduate School of Education Building (Moore Hall). This area comprises Room 145 on the main floor of the building, and Room 4 directly below it, which is reached by an interior stairway from Room 145. Each of these rooms is large and has been subdivided, in accordance with U.S.O.E. approval, into offices of various sizes. The total area is approximately 6,500 square feet.

In addition, CSE has a number of offices throughout the building assigned to faculty researchers. A large office on the third floor is assigned to the Higher Education Evaluation project and accommodates two researchers, four research assistants, and several clerical personnel.

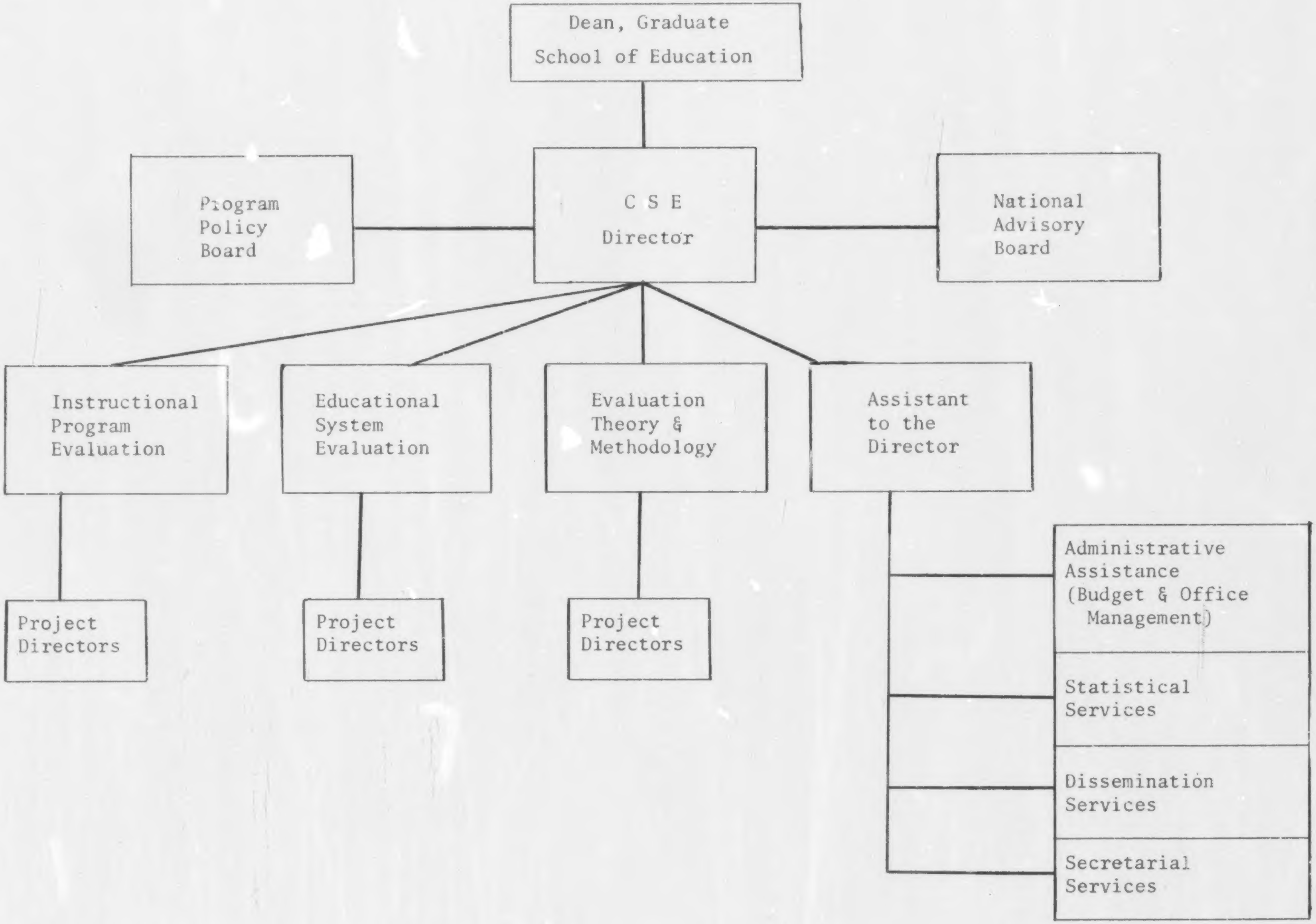
Although the University has the problem of providing space for all its activities, it has tried to be generous in assigning space to the Center. Nevertheless, the Center now has a problem of serious overcrowding.

MAJOR PROBLEMS

The major problem that CSE faces this year is the possible constriction of the budget to an amount between 90 and 100 percent of the last budget. This is considered a major constraint for several reasons: (1) The Center, as a relatively young organization, has been in its developmental stages. To cut its resources at this time would be to stifle its growth; (2) At a time when massive Federal expenditures have been provided for changing local school programs, accompanied with requirements that such programs be evaluated, there is an urgent need for improvement of evaluation procedures and techniques. When one considers the hundreds of millions of dollars that have been spent on evaluation by individual researchers, governmental agencies and private research organizations, it becomes difficult to imagine how the programmatic research thrust of this Center, funded at its present level, can be "all things to all people".

It is obvious that potential products of the Center, expressed as modified procedures for evaluating instructional programs, will pay for the cost of the Center's operations many-fold in the increased cost-efficiency of current instructional programs. What is needed, however, is an expenditure on research and development in the area of educational evaluation representing not a percentage increase in CSE's budget, but an order of magnitude increase.

ORGANIZATIONAL CHART - CENTER FOR THE STUDY OF EVALUATION


Dean, Graduate School of Education
Program Policy Board
C S E Director
National Advisory Board
Instructional Program Evaluation
Educational System Evaluation
Evaluation Theory & Methodology
Assistant to the Director
Project Directors
Project Directors
Project Directors
Administrative Assistance (Budget & Office Management)
Statistical Services
Dissemination Services
Secretarial Services

SUPPORTING SERVICES

STATISTICAL SERVICES

The Statistical Services Unit operates as a central service providing data processing and computer services to all programs and projects of the Center. This central service includes advice in planning a study, in data collection procedures, and in design of record forms, along with general guidance for assuring appropriate, accurate and efficient data processing procedures.

This unit is headed by George Simon. Dr. Simon is supported by programming and key punching personnel, including one full-time senior programmer, and three half-time junior programmers. There is a full-time supervisor of key punching and verifying operations, and one full-time key punch operator with several part-time operators.

HEAVY WORK LOAD

One need only look at the quantity of the work load for this unit to recognize the extent of its accomplishments during the past year. For example, key punching jobs were submitted from almost all programs and projects within the Center. From January through September, 1968, some 90,000 cards were punched and verified for approximately 50 different jobs.

The number of computer jobs has been increasing steadily. Over the past year, approximately 2,000 different jobs have been run. About one-fourth of these involved debugging of both old and new programs.

The great variety of jobs completed may be indicated by the following categories:

1) Data Screening and Descriptive Statistics -- including checks on maximum and minimum scores, checks for missing data, obtaining means and standard deviations, distributions and other such outputs.

2) Correlation Analyses -- including product-moment, rank difference, bi-serial, tetrachoric, etc.

3) Multivariate Analyses -- including multiple linear regression, stepwise regression, non-linear regression, principal components factor analyses, rotation of factors, etc.

4) Analyses of Variance and Covariance -- including simple and complex factorial designs, trend analyses and general linear hypothesis testing.

5) Test Analyses -- including item analyses, reliability, validity, etc.

6) Miscellaneous and Special Analyses -- including simulated (constructed) data, nonparametric tests,

conversions of data, and special outputs or report formats.

NEEDS ON THE INCREASE

Over the past year, requirements for statistical services support showed continual increase, and plans are being made to assure that future needs are adequately met. Among the improvements and programs undertaken during the year, the following may be noted:

1) Decrease in response time and turn-around time for getting jobs out.

2) Introduction of error control and error prevention procedures and general improvement in the efficiency of all operations.

3) Improvements in labeling and documentation of computer outputs and in control of data and procedures. Included here also is elimination of masses of printout sheets by designing outputs in a most useful format and minimizing number of sheets needed.

4) Improvements in old computer programs and preparation of new programs. A comprehensive effort is under way involving modular programming whereby it will be possible to put together sub-programs and sub-routines to meet new needs without having to write a new program every time a new requirement occurs. This effort also involves the writing of

programs that are more general and not limited to the particular application that generated them.

5) A program of orientation, communication, and training is under way not only for the central staff, but also for all users or potential users of central statistical services.

DISSEMINATION SERVICES

The need for a unique, effective program of dissemination for CSE is related both to the complexity and difficulty of its task of developing evaluation systems, and to the vital need for improved procedures for evaluating education today.

CSE can be considered but one agency involved in the study of evaluation, for everyone evaluating an educational program is in a sense making a contribution to the study of evaluation. There is a need to codify the knowledge relevant to evaluation, and to disseminate information among evaluation researchers. But the goal in dissemination is two-fold. We must keep informed not only the specialists actively involved in the Center, but also open to a more general public the nature and purpose of the Center's work.

EMPHASIS ON PUBLICATIONS

Publications are a major means of communicating the professional activities and scholarly contributions of the Center to the lay and professional publics, and to members of the Center itself. Publications, therefore, are vitally involved in both internal and external relations.

Internal relations involve communication among Center staff members, with the School of Education, and with the UCLA community. To promote vital internal communication, a monthly newsletter will be circulated to the Center staff, the Department of Education faculty, campus administrators (including department chairmen), and to other interested faculty members. The newsletter will include descriptions of the professional activites and on-going research of the Center staff, as well as periodic vignettes of the professional life of the different staff members. It might also include news about the non-academic staff members, thereby supplying much needed recognition of their contributions to the Center.

External relations primarily concern communication with the larger public outside the UCLA community, including the broad community of scholars as well as individual clients. Center publications should communicate in such a way that they will gain visibility for the Center and corresponding respect, interest, and use from an ever-increasing audience.

To attain these ends, the following procedures have been implemented:

1) An attractive brochure is being printed explaining the history and nature of the Center, the services it offers, and, briefly, representative professional activities of the members. The Center has been somewhat remiss, in the past, in this phase of its professional communication responsibility.

2) An annotated bibliography of official Center

publications will continue to be printed, revised as necessary, and circulated, along with the descriptive brochure. This bibliography notes which publications are available, provides a very brief synopsis of their content, and indicates how they can be obtained.

3) The _Evaluation Comment_ will be continued on a quarterly instead of a sporadic basis. This publication, initiated in January, 1968, has already gained national significance. Through it, work in evaluation study is exposed to a wide audience, and a forum for the exchange of ideas is provided. Hopefully, it will serve as a major vehicle for describing the research and the other important activities and ideas of the Center. An important and newly added feature of _The Comment_ will be summaries and interpretations of major research projects and subsequent reports, which can be read with interest and understanding by the educated layman as well as the scholar.

4) These publications will be distributed to as wide an audience as appropriate. This will include key administrators and superintendents of schools, research officers of the nation's colleges and universities, news services, selected members of such professional organizations as AERA, and other major public and private research centers and institutes with relevant interests. Dissemination Services will be responsible for these publications, and will pay particular attention to producing lively interpretive articles explaining the Center's activities.

5) The formal publications of CSE will continue to be published with a relatively small distribution, with the circulation population related primarily to professionals interested in the subject area of the paper. It will be an expectation that all papers will ultimately be available within the ERIC system, and that potential users will be referred to ERIC when CSE supplies of a particular paper have been depleted.

The formal paper series' are: 1) Working Papers, primarily conceptualizations or "think pieces" in an early state of development; 2) Occasional Reports, papers of the same type as above, but considered by the author to be in a somewhat finished stage; 3) Technical Reports, primarily the results of research studies; 4) Special reports of many types--bibliographies, instruction manuals, and reports of special interest to specific groups of educational practitioners.

PUBLICATIONS COMMITTEE

A permanent Publications Committee has been established to facilitate internal and external communications relations. This committee has a three-fold role: 1) As need arises, it will act as an advisory committee for the Director and the Program Planning Board on issues regarding publications; 2) It will act as a search committee by looking for ideas and research that merit publication, and should encourage staff members to develop such materials in a publishable form; 3) It will act as a screening committee to assure as much as possible the consistently high quality of all Center publications.

The committee is composed of at least four senior staff members who represent the various orientations and disciplines of the Center's professional staff. The Coordinator of Dissemination Services is a permanent member of this Committee. The other three members all serve one-year terms unless they are asked to, and are willing to, extend their terms. Appointments, or renewal of appointments, will be made formally on an annual basis.

POLICY AND PROCEDURES

The establishment of a publication policy is a necessity because the Center's primary form of communication will, of course, be through its formal publications. Therefore, these publications must have a valid purpose and be characterized by a high quality of substance and statement. The following procedures have been established to help assure the relevance and quality of Center publications:

1) The Publications Committee will encourage staff members to submit materials with publication potential to the Committee for review.

2) Staff members will be encouraged to duplicate and circulate their papers on a limited basis to colleagues for their interest and critical reactions.

It should be clearly evident, however, that these papers are not official publications of the Center. Unless the author has strong feelings against publication, he should automatically submit his papers for consideration by the Publication Committee.

3) It is assumed that final reports of approved CSE projects will be published. It is likely, however, that many project reports will need revision prior to formal publication.

4) All official publications of the Center, other than The Comment, must first be reviewed by the Publication Committee; no manuscript will be published without its recommendation. Exceptions are administrative documents such as the Annual Report, which would not normally be included in the Center's published bibliography.

5) Manuscripts will be evaluated by criteria employed by reputable professional journals or publishers. Materials will be judged for their substance first--that is for their originality, extent of unique contribution, internal logic, and substantiation of conclusions.

6) Where there is not a consensus of judgment within the committee, the evaluation of a qualified outside consultant will be sought. For the sake of impartiality the authorship of each manuscript will remain unknown to the consultant. Where a manuscript lies

beyond the areas of competency of a majority of the committee, again the judgment of a qualified outside consultant will be sought. Members of the committee will not review their own manuscripts. The committee will make every effort to indicate revisions that would render manuscripts acceptable for publication rather than reject them outright, with the exception of papers considered unredeemable. To the extent possible, the committee will make suggestions having to do with major problems of substance and expression such as techniques of data analysis or organization of the manuscript. All accepted manuscripts will be submitted to the Office of Dissemination Services for editing, at which time the author may frequently have to be involved in subsequent revisions. At all times it is the author's responsibility to communicate as clearly and as professionally as possible.

7) The senior author of a Center publication, other than contributors of The Comment, must be a junior or senior staff member of the Center, except for extraordinary circumstances determined by the Director. Junior staff members, such as research assistants, are to be encouraged to submit materials. The publication of dissertations done in conjunction with the Center is also to be encouraged. These materials, however, must be reviewed in the same manner as all other manuscripts, and it is expected that most dissertations will require considerable revision for publication. It is to be clearly understood that the

Publications Committee will make no judgments about the worth or acceptability of a dissertation as a dissertation, but only as a Center publication.

8) It shall be a responsibility of the Publications Committee to advise on such matters as titles, cover design, publication format, and charges to be made for publications.

9) Dual publication is to be encouraged. Center publications should not preclude subsequent journal or commercial publication. Articles resulting from Center activity should note this fact. The Center, in turn, will purchase reprints of these articles and bind them for distribution. Articles published elsewhere after Center publication should have this noted on the Center's materials.

10) A wide variety of publications are to be encouraged: study of evaluation gives rise to many issues, techniques, philosophies, theories, and contexts as well as much data. What one staff member may see as tangential to the Center's objectives may be viewed as essential to these objectives by another. There must, however, be some real, observable relationship between each of the Center's publications and its objectives. This relationship must be stated explicitly in some form in each manuscript. It shall be a responsibility of the Publications Committee to help discern this relationship if it is not clear, to suggest modifications accordingly, or to reject manuscripts which are not relevant to the Center's

objectives.

To carry on this expanded program the dissemination and editorial services of the Center have been restructured. Sherman Pearl has been employed as Coordinator of Dissemination Services for CSE. Mr. Pearl was previously employed in Washington, D. C. as a writer/editor by the U. S. Department of Interior and the Office of Economic Opportunity.

CLERICAL SERVICES

CSE has felt it beneficial to provide a high degree of centralization of much of the secretarial and clerical services of the Center.

While some major projects have individual secretarial assistance, it is felt that greater efficiency can be achieved not only in cost savings, but also in quality of work produced, by maintaining a central services unit for these functions.

The centralized secretarial unit not only provides clerical assistance to the projects, but also prepares all publications, including journal articles, manuscripts, interim and final reports, and other material. The service, headed by Mrs. Joyce McNamara, works closely with the Office of Dissemination Services in the production of Center publications.

CSE STAFF
AND
NATIONAL ADVISORS

THE CENTER'S STAFF

MARVIN C. ALKIN, the new CSE director, is an Associate Professor of Educational Administration. As a member of the Center's founding committee, he helped write the original research proposal. Dr. Alkin is an alumnus of the Stanford University School of Education. His papers include such subjects as mathematical models for school district evaluation, data accessibility in school district research, economy of scale in education, and cost-effectiveness evaluation of education. Dr. Alkin has directed a project on developing a simulated evaluation exercise for use in training educational evaluators.

EVA BAKER, an Assistant Professor of Curriculum and Instruction and of Programmed Instruction at UCLA, is currently conducting research in the areas of teacher behavior, instructional variables, and pupil evaluation. She holds a Master's and Ed.D., from UCLA's Graduate School of Education in the areas of programmed instruction and research design. In the past she has worked at the Southwest Regional Laboratory for Educational Research and Development on a project to determine objectives for instructional procedures. Her other research interests include verification of effects of instructional variables, and teacher evaluation.

MARY M. BENTZEN, a specialist in the sociology of education, has been the Center's project coordinator for LAMMP--a model mathematics project undertaken in Los Angeles schools. Her background includes group work for Probation Department, teacher supervision, and research, teaching and administration in elementary school. She has also worked extensively in production of educational films and videotapes. These deal with such subjects as school experiences of disadvantaged children, laboratory schools, group counseling with parents, principles of nongrading and team teaching, and foreign language teaching methodology. Among her present research interests are studies of educational change and organizational variables.

CLARENCE BRADFORD, an Assistant Professor at UCLA, has major interests in the fields of analysis of educational research data, experimental design in educational research, and analysis of survey data in education. He is currently working on the interpretation of techniques of canonical regression as they apply to behavioral systems. Another current project is the construction of models of sample-free item statistics. Dr. Bradford has been an associate director of test construction projects at the University of Chicago, a project director at Science Research Associates, and an Assistant Professor at Washington University in St. Louis. He holds an M.A. in philosophy and a Ph.D. in education from the University of Chicago.

NORMA FESHBACH, a clinical and developmental research psychologist, is an Assistant Professor of Educational Psychology. She heads the current U.S.O.E. projects on cross-cultural comparison of teaching styles in young children, and on construction of an Attitude Toward School Instrument. Prior to her doctorate work in psychology at the University of Pennsylvania she worked as a teacher and lecturer, researcher and psychology intern in eastern hospitals, clinics, and colleges. She later continued her teaching and research at Stanford, University of California at Berkeley, University of Colorado, and UCLA. Her writings have dealt with such subjects as teacher's reinforcement style, aggression, empathy, and cognitive development.

C. WAYNE GORDON, Associate Dean of UCLA's Graduate School of Education, is currently involved in a wide range of research projects. His studies vary from a demographic analysis of native Alaskan village populations, to educational achievement and aspirations of metropolitan area Mexican-American youth. Relevant to his relationship with the Center are his studies on social systems in the high school and on dimensions of teacher leadership. Dean Gordon's 30-year experience in the field of education has included teaching assignments on every level. Specializing in sociology and the sociology of education, he directed the University of California's Study Abroad Center in Hong Kong in 1965-66. He was also External Examiner for the Department of Education at the University of Hong Kong.

THEODORE R. HUSEK, a specialist in psychological measurement and mathematical statistics, heads two CSE projects studying criterion variables. Dr. Husek has taught statistics and measurement in the psychology departments at the University of Illinois (where he received his Ph.D.), Hawaii, and Washington. In the Graduate School of Education at UCLA he is a Professor specializing in underlying theory and special problems of measurement in education. Consultant editor to the American Journal of Educational Research, and consultant to the research and development unit of Los Angeles City Schools, his major topics of publication have been test theory and construction, social psychology, and measurement of anxiety.

MARGARET H. JONES, a research psychologist in UCLA's departments of Education and Psychology, has undertaken R & D Center work studying psycholinguistics, perception, and cognitive development in order to determine how these affect comprehension of language. She earned her B.A. at Vassar, M.A. at Hobart College, and Ph.D. at UCLA, Dr. Jones was formerly an associate director of the Engineering Qualifying Examination Project, and was also affiliated with the University of Alabama and State University of Washington before coming to UCLA to lecture in education. Author of two books and 35 articles, she is presently a co-director of the CSE project on Contextual Constraints in the Language of the Child.

STEPHEN P. KLEIN comes to CSE from Princeton, New Jersey, where he was a research psychologist and project director with the Educational Testing Service. He is director of a newly formed project for developing a broad school evaluation system. After undergraduate work at Tufts University, Dr. Klein received his Ph.D. in Industrial Psychology from Purdue in 1965. He has published extensively in journals of education and psychology on the subjects of creativity, academic prediction, test evaluation, and analysis of graduate instructional programs. Dr. Klein has also co-authored a series of articles on the measurement and prediction of artistic creativity and aesthetic judgement.

ERICK L. LINDMAN a former director of CSE and Acting Dean of the UCLA School of Education, is currently involved in Center-sponsored projects aimed at developing a model for evaluating secondary school systems and a management system for elementary schools. He has long been active in the field of state school administration and finance, serving as consultant for educational studies in 11 states. He has worked in government as chief of the school administration branch of U.S.O.E. and as chairman of the committee on educational finance of the National Education Association. Dr. Lindman, who received his M.A. and Ph.D. in Education from the University of Washington, was also a former Deputy Superintendent of Public Instruction for the State of Washington.

DAVID O'SHEA, Assistant Professor of Education at UCLA, is participating in the Center's project to develop a test kit for elementary school evaluation. Having received a Bachelor's degree in Sociology from the University of Chicago, Professor O'Shea's graduate research has focused on the politics of local school districts. Ecological factors governing resource distribution, and socio-economic factors influencing learning were some of the areas he studied. In the past, Professor O'Shea has conducted research on society and education, and on standardization of grading practices. He currently teaches graduate courses in the Sociology of Education and Educational Organization.

C. ROBERT PACE has directed Center projects on the development of contextual and criterion measures for higher education. His major research interests in higher education are evaluation and measurement, college environments, and follow-up studies of college graduates. Educated at De Pauw University, with a M.A. and Ph.D. in Educational Psychology from the University of Minnesota, Dr. Pace has been, since 1961, a Professor of Higher Education at UCLA. Previous affiliations were with the Evaluation Center and the Psychology Department at Syracuse University, the Personnel Research Section of the Bureau of Naval Personnel, and the American Council on Education.

W. JAMES POPHAM, an Associate Professor at UCLA, is deeply involved in various aspects of educational technology. He is currently directing CSE's Instructional Objectives Exchange. His most recent research has been in the evaluation of programmed instructional materials for Teacher Education. In the past five years he has also conducted U.S.O.E.-sponsored projects to develop performance tests of instructor competence. An alumnus of the University of Portland, he taught at San Francisco State College and at Indiana University, where he earned his Ed.D. in Secondary Education. Dr. Popham recently completed a book, *Educational Statistics: Use and Interpretation*, published by Harper and Row.

AUDREY SCHWARTZ, a lecturer in the Departments of Education and Sociology at UCLA, is currently studying the discrepancy between teacher assigned marks and objective test scores. The study seeks to clarify the sources and effects of the discrepancy and their relation to the functioning or dysfunctioning of the school system. Dr. Schwartz holds Bachelor's and Master's degrees from the University of Pennsylvania and a doctorate in Sociology of Education from UCLA. In the past, as a research educationist for CSE, she has researched, written and lectured on the values, aspirations, and achievements of Mexican-American youth.

GEORGE B. SIMON heads the Statistical Services Unit at CSE. He has also been director of the Center's program for developing an Elementary School Evaluation System. Holding an Ed.D. in Research and Evaluation from Harvard University's School of Education, Dr. Simon has an extensive background in test development, industrial psychology, data processing, information systems, and simulation and criteria research. His 36 technical publications deal with many of these areas. A member of the Center staff since 1967, Dr. Simon is a full-time research educationist.

RODNEY W. SKAGER has been a program director with CSE since its inception in 1966. His recent projects have centered on the use of cognitive development measures in evaluation of instructional programs. A graduate of the University of Redlands, he holds an M.A. and Ph.D. in psychology from UCLA. An Assistant Professor of Education and a research psychologist, his research interests have centered on the evaluation of instructional programs, the measurement of cognitive development, and the function of educational research in the process of making decisions in education. The latter area is the subject of a forthcoming book.

GARTH SORENSON, Professor of Educational Psychology and Counseling Theory in UCLA's Graduate School of Education, heads the CSE project on research in counseling. In his 15 years at UCLA, Dr. Sorenson, who holds advanced degrees in Sociology and Educational Psychology from the University of Utah, has conducted research in teacher education, counselor training, and counselor effectiveness. He has written reports for the R & D Center on development of an instructional model for counseling, and on the matching process in teacher appraisal. He also serves as a consultant to various public and private schools in the Los Angeles and San Diego areas.

JAMES W. TRENT serves the R & D Center in projects relating to higher education, a specialization in which he is an Assistant Professor at UCLA. Specifically, he will be working to develop criterion and contextual measures for this field. He comes to the Center from the U.S.O.E. R & D Center at Berkeley. Dr. Trent has written extensively on psycho-sociological studies of high school graduates, including studies of characteristics associated with various patterns of college attendance, attitude changes of college and noncollege peer groups, the intellectual development of college students of different religious subcultures, and studies of activism and apathy among college students. Since receiving his Ph.D. at Berkeley in 1964, Dr. Trent has served widely as a consultant on student development, graduate education planning, social science research in relation to counseling, research in multiple-media teaching techniques, self study projects, and junior college programs for cultural minorities.

LOUISE TYLER, an Associate Professor of Education at UCLA, is actively involved in program planning for CSE. Her central interests lie in the field of curriculum construction and evaluation. She draws upon the theoretical framework of psychology and related disciplines for investigations in the field. Her recent research interests have included the formulation of criteria for evaluating curriculum and instructional materials, and the formulation of objectives for an instructional program based upon psychoanalysis. She has been director of a CSE project for assessing ego strength. Dr. Tyler teaches courses in Principles of Curriculum and Instruction, as well as Evaluation.

CARL WEINBERG, who teaches Counseling and Sociology of Education at UCLA, has recently completed a CSE study project on the alienation of teachers. He is currently working on the preliminary draft of an instrument to measure student alienation. He has completed other studies on student deviance, race relations in schools, student leadership, and differential treatment of students as based upon economic and racial characteristics. Professor Weinberg, who earned his Ed.D. at UCLA has published a book entitled Social Foundations of Educational Guidance and is currently preparing books on research for beginning students in education, and education and social problems.

MERLIN C. WITTROCK, an educational psychologist, is co-director of CSE. He is also chairman of the CSE Program Policy Board, and he is directing a major program on the evaluation of classroom interaction. Since completing his Ph.D. at the University of Illinois in 1960, he has been teaching and doing research on children's learning and instruction at UCLA. He is a leading contributor to research in these areas. He is currently serving as editor-in-chief of a seven-volume work for the American Educational Research Association, Readings in Educational Research. In 1967-68 he was a fellow at the Center for Advanced Study in the Behavioral Sciences at Stanford, California. He is currently Chairman of the Conference of Educational R & D Center Directors.

CSE NATIONAL ADVISORY BOARD

The UCLA R & D Center is aided by a National Advisory Board of distinguished professors. The members are:

DR. BENJAMIN S. BLOOM, Professor of Education at the University of Chicago;

DR. JOHN O. DARLEY, Chairman of the Department of Psychology, University of Minnesota;

DR. BURKHARDT HOLZNER, Chairman of the Department of Sociology, University of Pittsburgh;

DR. H. THOMAS JAMES, Dean of the School of Education at Stanford University.

A SUMMARY OF PROGRESS

A SUMMARY OF PROGRESS

For the past year, the Center has conducted a wide range of research studies dealing with a number of elements of total evaluation systems. The studies dealt with instructional, contextual, and criterion variables, and CSE programs were organized in those three areas.

Although CSE has recently reorganized its studies for a more integrated, comprehensive approach, this past work produced valuable contributions to the study of evaluation, and laid the groundwork for the Center's new programs. The anticipated products of continuing or new projects, in terms of their potential contributions to education, are also great.

ADVANCES BY TERMINATING PROJECTS

A brief summary of the findings of a few past projects will indicate the breadth of involvement of the Center and the contributions it has already made to evaluation.

RESEARCH ON COGNITIVE SKILLS: This project required the development or adaptation of a variety of measures of level of cognitive development for use in the evaluation of an instructional program. If preliminary results are verified in analyses presently underway, the project will represent a significant first step toward the utilization of such measures in other types of evaluation research.

In contrast to standardized measures of aptitude and achievement, information about the cognitive skills of children provides a picture of intellectual development in terms of where the child is in the growth process. That is, such measures do not merely indicate how many items a child passes in comparison with a norm group, but rather were designed to shed light on the types of skills and logical processes that are available for the solution of many types of problems.

CSE is attempting to determine whether such skills were related to gains in achievement in mathematics over the school year. If such is the case, evaluation data will be rendered considerably more informative by making it possible to determine whether programs are more or less appropriate for children at certain levels of cognitive development.

Such information provides a type of evaluation data often praised but rarely obtained. Instead of merely reporting whether a program is effective or ineffective for children in general, CSE holds that the evaluator should attempt to take into account the relation between characteristics of students and the outcomes of instruction. The utilization of cognitive measures appears to be the most logical place to begin.

STUDY OF INSTRUCTIONAL VARIABLES: The five projects within this program of evaluation of the learning, retention, and transfer from instruction of elementary school children focused on the effect of memory and perceptual traits on classroom learning, and on the effect of organization of the instructional stimuli on the results of instruction.

Important advances were made in (1) assessing the importance of individual differences in memory abilities as they affect classroom instruction; (2) exploring the relationships among units, classes, and systems of knowledge, and whether this system can be used as a hierarchy for the purpose of classroom materials; (3) studying the effect of generality of rules--whether general rules or specific are superior to each other and to examples in promoting transfer of learning; (4) measuring individual differences in how, and the extent which, elementary school children can be taught to group perceptual units; (5) investigating the effects of organization upon learning to determine whether it is important to index the degree of hierarchical organization of instructional materials.

MINORITY YOUTH EDUCATION: The improvement of education of minority group youth is one of the most crucial problems facing the nation's educators. CSE's work in this area has provided valuable information which can point the way toward solutions.

Elaborating upon the findings of the Coleman Report, CSE conducted local studies dealing with the educational achievements of Mexican-American students. A wide range of variables was investigated, which provided valuable information on those variables most strongly influencing the educational outcomes of minority students.

CSE research is timely for educational systems in the process of forming policy on the education of minority students. The Coleman Report showed several deficiencies in the schooling of Mexican-Americans. Center studies provide evidence that the educational environment must be improved for these students.

These findings have significant implications affecting not only local Mexican-American youth, but the education of urban youth nationally.

POTENTIALS OF CONTINUING PROJECTS

Examples of CSE projects being continued or expanded will suggest the importance of their outcomes and potential benefits to improving the evaluation of American education.

MATRIX SAMPLING: If the multiple outcomes of educational programs are to be assessed, techniques for measuring a broad variety of outcomes are essential. Simply having many types of tests does not solve the problem if cost or other practical limitations make it impossible to obtain scores on the tests. CSE research on matrix sampling may offer a solution.

For many educational evaluation problems, matrix sampling provides the only method for collecting adequate data. It is not limited to achievement or cognitive tests, nor to binary items. Since it is a general procedure, it need not be limited to pencil and paper items, nor must its use be limited to one test.

Matrix sampling involves simultaneous, random sampling of both students and items. Where both matrix sampling and simple sampling of examinees are possible, the former is generally the more efficient for the estimation of the population mean parameter.

The determination and use of appropriate sample sizes for students and items may lead to the ability to use limited school district resources more efficiently. Using matrix sampling, it is anticipated that either the same amounts of student test time (with attendant financial resource requirements) will provide valid data on more outcome dimensions, or less test time will provide outcome data similar to that produced without sampling.

ELEMENTARY SCHOOL EVALUATION: The elementary school project is the initial stage of what is intended to be a comprehensive school information system. This system will serve both as a body of carefully collected data showing the power and/or weaknesses of alternative conceptualizations and strategies for evaluation of the schools, and as an environment for researchers to explore new dimensions and procedures.

The project is evolutionary rather than static. It is an organization of the changing thrusts which may be vital to an understanding of the functions and products of our schools. As the system grows more comprehensive and precise, its direction of effort will undoubtedly change. It is intended to serve the administrator who asks for carefully documented evidence about some kinds of information or the value of particular evaluation instruments, the innovator as a source of information to explore how new inputs to the schools may affect the product, and the researcher who seeks to examine new relationships or to simulate new conditions.

The eventual product of this school evaluation system will be a body of evidence covering the full breadth of variables relating to both the inputs and the outputs of the schools. It will above all serve as a reliable source of information which school administrators and teachers can use for making changes that upgrade the results of their professional efforts.

HIGHER EDUCATION: The Center's activities in evaluation of higher education have developed from two initial beliefs: (1) that the diversity of higher education in the United States requires a corresponding diversity in the criteria by which institutions are evaluated; (2) that major differences in institutional environments need to be identified and measured more effectively so that clearer relationships between treatments and outcomes can be established.

The national field study of higher education embodies CSE work in expanding the range of criterion and conceptual variables. The range of criterion measures is really the crucial difference between this and previous studies. It is one thing to identify personal dispositions, educational experiences, and environmental conditions that are predictive of our criterion--such as those concerning one's education in graduate school; but it is quite another thing to discover whether the dispositions, experiences, and conditions positively associated with going to graduate school may be negatively associated with involvements in community service, politics art, music, literature and drama.

It is the concern with different outcomes, each of which is important and relevant to education and the larger society, that gives a new level of complexity to this CSE study. If a national study of higher education does not include a broad range of possible outcomes and potential contributions, then it is not, and should not, be interpreted as an evaluation.

This level of complexity has demanded development of a concept of evaluation that differs considerably from past views, as follows:

1. It begins with the question, "What are possible consequences?" rather than with the more limiting question, "What are the specific objectives?"

Its style of inquiry is more aptly characterized by the word "exploration" than by the words "control" and "focus".

3. It sees the role of the evaluator as a social scientist rather than as a teacher, or reformer, or staff officer to the practitioner.

4. Its purpose is to provide more complex bases for informed judgment.

What this approach means, in terms of its potential impact on higher education, is seen by the Center as follows:

1. Influencing the way people think about evaluation, so that their work may be more relevant.

2. Influencing the practice of evaluation in institutional research offices across the country, and by other investigators by making available to them new measuring instruments of demonstrated reliability and importance.

3. Influencing national thinking about higher education by presenting from our data conclusions and judgments that are more balanced than many that are commonly expressed.

4. Influencing the effectiveness of institutions by demonstrating the connections between various practices and outcomes.

SIMULATED EVALUATION EXERCISE: This exercise will allow participants to act as evaluators and to construct evaluation designs which are capable of modification to meet the demands and problems of actual field conditions. A team will evaluate experimental mathematics for underachieving junior high school students. Participants will submit the best possible design to deal with field contingencies.

The completed exercise will consist of bibliography, instructor guide and material supplement, mock final report, simple version of the exercise, and a "how to do it kit".

The entire package will provide trainers of potential evaluators with a basic evaluation unit whose elements can be combined to meet various training needs. The prospective evaluator is thus introduced to the problems and frustrations that occur in field study, and learns the prerequisites for design construction and flexibility that will enable him to be most effective as an evaluator.

It is anticipated that the total evaluation exercise will be of great value as a part of formal as well as in-service training programs for those who have a responsibility for educational evaluation. CSE anticipates great value from this project in improving the quality of educational evaluation in field settings.

NEW PROJECTS MOST PROMISING

It is in the realm of the Center's newly initiated projects that the potential of the Center is most fully seen. Two new Center projects have especially great implications for changing educational evaluation.

INSTRUCTIONAL OBJECTIVES EXCHANGE AND MEASUREMENT SYSTEM: CSE has initiated a service which will use a wide range of categorization rubrics to supply initial input for providing schools with objectives and test measures of as many areas as they may need.

These projects will greatly help teachers obtain feedback on the extent to which instructional objectives have been met, whether the goal is evaluating the learner or the instructional program. In final form, the program would enable a teacher to select from a file of objectives for a specific subject area those are which are most appropriate for her classroom. The teacher would then order a test or series of items covering the relevant skills; the instrument will be produced almost immediately by selection of the test items from a bank of items coded into a computer by objective.

After testing, results would be scored, providing input for generating reports to the teacher. Depending upon whether the teacher were interested in the diagnosis of individual progress or in identifying defects in the instructional program, scores could be reported either for individual students or for the group as a whole in each skill area.

The major contribution of these projects is that they will provide a means for generating and using content relevant tests on a large scale basis. Thus, classroom instruction and school systems can be evaluated in terms of their own objectives.

BENEFITS TO AMERICAN EDUCATION

Discussion of these past and present projects offers some measure of the involvement, contributions, and potential benefits of CSE to American education and evaluation. Projects selected for further development are those which most fully utilize the total resources of the Center's program structure and which focus upon the needs and problems of educators and evaluators throughout the country.

These projects are pursuing the goal of continual and expanding publications of their findings. Dissemination policy is to identify the audiences for whom the results of specific programs are most relevant, and to furnish them with significant and meaningful reports and papers.

These projects and policies promise to bring CSE into an increasingly vital role as the nation's principal investigator of long-deferred evaluation problems.

BIBLIOGRAPHY

BIBLIOGRAPHY

(CSE Publications Currently in Print)

CSE REPORT	TITLE	AUTHOR
Working Paper #1 November, 1966	Categories of Cognitive Skills	Harriet Foster
Working Paper #2 December, 1966	Experiment in Research on Evaluation of Instruction	M. C. Wittrock
Working Paper #3 September, 1968	Dimensions of the School Budget	Erick L. Lindman
Working Paper #4 August, 1967	Towards an Evaluation Model: A Systems Approach	Marvin C. Alkin
Working Paper #5 March, 1968	Net-Shift Analysis for Comparing Distributions of Test Scores	Erick L. Lindman
Working Paper #6 June, 1968	A Three-Dimensional Program Account Classification System for Public Schools	Erick L. Lindman
Occasional Report #1 February, 1967	Cloze Readability Procedure	John R. Bormuth
Occasional Report #2 February, 1967	Item Sampling in Educational Research	T[illegible] R. Husek and K[illegible]. Sirotnik
Occasional Report #3 April, 1967	The Implications and Use of Cloze Procedure in the Evaluation of Instructional Programs	John R. Bormuth
Occasional Report #4 April, 1967	Teacher Appraisal: A Matching Process	Garth Sorenson and Cecily F. Gross
Occasional Report #5 April, 1967	Cognitive Structures and Educational Evaluation	R. W. Skager and L. A. Broadbent
Occasional Report #6 April, 1967	Toward an Instructional Model for Counseling	Garth Sorenson

Bibliography 2

CSE REPORT	TITLE	AUTHOR
*Occasional Report #9 October, 1968	Toward a Theory of Testing Which Includes Measurement-Evaluation-Assessment	Benjamin S. Bloom
*Occasional Report #10 October, 1968	Evaluation as a Main Aim of Science: Comments on Professor Bloom's Paper (Occasional Report #9)	Michael Scriven
*Occasional Report #11 October, 1968	Comments on Professor Bloom's Paper (Occasional Report #9)	Gene V. Glass
**Occasional Report #12 October, 1968	Comments on Professor Bloom's Paper (Occasional Report #9)	J. P. Guilford
*Occasional Report #13 October, 1968	Theory of Evaluation of Instruction	Robert Glaser
*Occasional Report #14 October, 1968	Comments on Professor Glaser's Paper (Occasional Report #13)	Robert Stake
*Occasional Report #15 October, 1968	Comments on Professor Glaser's Paper (Occasional Report #13)	Arthur Lumsdaine
*Occasional Report #16 October, 1968	Instructional Variables and Learning Outcomes	Robert Gagné
*Occasional Report #17 October, 1968	Comments on Professor Gagné's Paper (Occasional Report #16)	Richard Anderson
*Occasional Report #18 October, 1968	Comments on Professor Gagné's Paper (Occasional Report #16)	Leo Postman
*Occasional Report #19 October, 1968	The Cracked Cake of Educational Customs and Emerging Issues in Evaluation	Dan Lortie
*Occasional Report #20 October, 1968	Comments on Professor Lortie's Paper (Occasional Report #19)	C. Wayne Gordon
*Occasional Report #21 October, 1968	Comments on Professor Lortie's Paper (Occasional Report #19)	N. L. Gage

Bibliography 3

CSE REPORT	TITLE	AUTHOR
Technical Report #1 February, 1967	Input-Output Relationships in a Sample of California Public Junior Colleges	Marvin C. Alkin and Vernon L. Hendrix
Technical Report #2 February, 1967	Variations in Teachers' Reinforcement Style and Imitative Behavior of Children Differing in Personality Characteristics and Social Background	Norma D. Feshbach
Technical Report #3 December, 1967	Psychophysiological Correlates of Female Teacher Behavior and Emotional Stability: A Seven-Year Longitudinal Investigation	William H. Lucio, Marion A. Wenger and Thomas D. Cullen
Technical Report #4 April, 1968	The Measurement of Campus and Student Morale	Lora Robinson and Richard Seligman
Special Report #1 June, 1968	CSE Simulated Evaluation Exercise	Marvin C. Alkin, Maxine Bentzen, J. Eugene Grigsby III, Stuart Johnson and Rita Johnson

* - These publications are in the process of being printed.
Abstracts are also being prepared.

** - In print; abstract being prepared